



HAL
open science

Nouvelle approche basée sur la dynamique moléculaire et la RMN pour l'étude et l'optimisation de nouveaux inhibiteurs de peroxyrédoxines humaines, impliquées dans les chocs ischémiques

Laura Troussicot

► **To cite this version:**

Laura Troussicot. Nouvelle approche basée sur la dynamique moléculaire et la RMN pour l'étude et l'optimisation de nouveaux inhibiteurs de peroxyrédoxines humaines, impliquées dans les chocs ischémiques. Chimie analytique. Université de Lyon, 2017. Français. NNT : 2017LYSE1071 . tel-01559433

HAL Id: tel-01559433

<https://theses.hal.science/tel-01559433>

Submitted on 10 Jul 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



N° d'ordre NNT : 2017LYE1071

THESE de DOCTORAT DE L'UNIVERSITE DE LYON

Opérée au sein de
L'Université Claude Bernard Lyon 1

Ecole doctorale de chimie ED 206

Spécialité de doctorat : Chimie

Discipline : Analyses biophysiques et chimiques

Soutenue publiquement le 05/05/2017, par :

Laura TROUSSICOT

Nouvelle approche basée sur la dynamique moléculaire et la RMN pour l'étude et l'optimisation de nouveaux inhibiteurs de peroxyrédoxines humaines, impliquées dans les chocs ischémiques

Devant le jury composé de :

KNOOPS, Bernard	Professeur	Université Catholique de Louvain	Rapporteur
TAREK, Mounir	DR	Université de Lorraine	Rapporteur
MIELE, Adriana	Professeur	Université Lyon 1	Examineur
SIBILLE, Nathalie	CR1	CBS (Montpellier)	Examineur
ROCHE, Didier	VP	Edelris (Lyon)	Examineur
LANCELIN, Jean-Marc	Professeur	Université Lyon 1	Directeur de thèse
GUILLIERE, Florence	MCF	Université Lyon 1	Membre invité

UNIVERSITE CLAUDE BERNARD - LYON 1

Président de l'Université

M. le Professeur Frédéric FLEURY

Vice-président du Conseil d'Administration

M. le Professeur Hamda BEN HADID

Vice-président du Conseil d'Administration

M. le Professeur Didier REVEL

Vice-président du Conseil Formation et Vie Universitaire

M. le Professeur Philippe CHEVALIER

Vice-président de la Commission Recherche

M. Fabrice VALLÉE

Directrice Générale des Services

Mme Dominique MARCHAND

COMPOSANTES SANTE

Faculté de Médecine Lyon Est – Claude Bernard

Directeur : M. le Professeur G.RODE

Faculté de Médecine et de Maïeutique Lyon Sud – Charles Mérieux

Directeur : Mme la Professeure C. BURILLON

Faculté d'Odontologie

Directeur : M. le Professeur D. BOURGEOIS

Institut des Sciences Pharmaceutiques et Biologiques

Directeur : Mme la Professeure C. VINCIGUERRA

Institut des Sciences et Techniques de la Réadaptation

Directeur : M. X. PERROT

Département de formation et Centre de Recherche en Biologie Humaine

Directeur : Mme. la Professeure A-M. SCHOTT

COMPOSANTES ET DEPARTEMENTS DE SCIENCES ET TECHNOLOGIE

Faculté des Sciences et Technologies

Directeur : M. F. DE MARCHI

Département Biologie

Directeur : M. le Professeur F. THEVENARD

Département Chimie Biochimie

Directeur : Mme C. FELIX

Département GEP

Directeur : M. H. HAMMOURI

Département Informatique

Directeur : M. le Professeur S. AKKOUCHE

Département Mathématiques

Directeur : M. le Professeur G. TOMANOV

Département Mécanique

Directeur : M. le Professeur H. BEN HADID

Département Physique

Directeur : M. J-C PLENET

UFR Sciences et Techniques des Activités Physiques et Sportives

Directeur : M. Y.VANPOULLE

Observatoire des Sciences de l'Univers de Lyon

Directeur : M. B. GUIDERDONI

Polytech Lyon

Directeur : M. le Professeur E.PERRIN

Ecole Supérieure de Chimie Physique Electronique

Directeur : M. G. PIGNAULT

Institut Universitaire de Technologie de Lyon 1

Directeur : M. le Professeur C. VITON

Ecole Supérieure du Professorat et de l'Education

Directeur : M. le Professeur A. MOUGNIOTTE

Institut de Science Financière et d'Assurances

Directeur : M. N. LEBOISNE

RESUME

Nouvelle approche basée sur la dynamique moléculaire et la RMN pour l'étude et l'optimisation de nouveaux inhibiteurs de peroxyrédoxines humaines impliquées dans les chocs ischémiques

La description des interactions protéine-ligand est d'une importance capitale pour la compréhension de phénomènes biologiques de toutes sortes, et dans le procédé de conception de nouvelles molécules bioactives. Avec l'avènement de nouveaux moyens de calculs numériques, la modélisation des équilibres d'interaction utilisant la dynamique moléculaire est de plus en plus utilisée pour l'étude au niveau microscopique permettant ainsi de guider la conception et l'optimisation de nouveaux inhibiteurs biologiques.

Cette thèse est centrée sur l'utilisation de la funnel-métadynamique pour la prédiction d'affinités protéine-ligand et la description des interactions, en corrélation avec des données expérimentales de RMN et de cinétique enzymatique. Sur le modèle biologique des peroxyrédoxines humaines impliquées dans la cascade inflammatoire survenant après un choc ischémique, l'interaction et l'inhibition de composés catéchols ont été étudiées. Les données obtenues par dynamique sur l'interaction de dérivés catéchols ont été utilisées pour guider l'optimisation de cette molécule fragment pour une meilleure affinité et inhibition des peroxyrédoxines humaines. Nos résultats ont montré que la funnel-métadynamique, en plus de permettre la prédiction des affinités protéine-ligand, donnait une description réaliste de l'interaction pouvant mener à l'identification et l'optimisation de nouvelles molécules bioactives dont le potentiel inhibiteur a pu être examiné par cinétique enzymatique. Cette recherche fournit un aperçu des possibilités offertes par les nouvelles méthodes numériques, et leur application en chimie médicinale par exemple.

Mots clés : *Funnel-métadynamique, Interactions, RMN, Catéchol, Peroxyrédoxines humaines, Inhibition*

ABSTRACT

New approach based on molecular dynamics and NMR to study and optimize new inhibitors of human peroxiredoxins involved in ischemic strokes

Description of protein-ligand interactions is crucial for the understanding of all biological phenomena, and the drug design process. Thanks to new developments in computational devices, simulation of interaction equilibria using molecular dynamics are becoming state-of-the-art approach for the microscopic study of these molecular interactions. These new methods guide the conception and the optimization of new biological inhibitors.

This project is focused on the use of funnel metadynamics for the prediction of protein-ligand affinities, and the description of interactions. Data obtained by metadynamics are correlated with experimental data from NMR and enzymatic assays. On the biological model of human peroxiredoxins, involved in the post ischemic inflammation cascade, interaction and inhibition of catechols derivatives have been studied. Data obtained from simulations have been used in the optimization process of the catechol fragment, to reach a better affinity and inhibition of human peroxiredoxins.

Our results have shown that funnel-metadynamics allows the prediction of protein-ligand affinities, and the realistic description of the interaction, that lead to identify and optimize new bioactive molecules. Their inhibitory strengths and mechanisms have been examined using enzymatic assays.

Our research gives an overview of the possibilities brought by new numerical approaches, and their application in medicinal chemistry for example.

Key words : *Funnel metadynamics, Interactions, NMR, Catechols, Human Peroxiredoxins, Inhibition*

REMERCIEMENTS

La rédaction de ce mémoire me donne l'opportunité de remercier toutes les personnes qui ont contribué de près ou de loin à l'accomplissement de cette thèse, et m'ont soutenue durant ces 3 ans.

Je tiens tout d'abord à remercier les membres du jury qui m'ont fait l'honneur de d'examiner mes travaux de recherche : les deux rapporteurs Pr. Bernard KNOOPS et Dr. Mounir TAREK, ainsi que les examinateurs Pr. Adriana MIELE, Dr. Nathalie SIBILLE et Dr. Didier ROCHE. Je remercie également les membres de mon comité de thèse : Pr. Christophe MORELL et Dr. Sébastien VIDAL pour leurs conseils et pour les discussions enrichissantes que nous avons pu avoir sur ce projet.

Cette première expérience professionnelle dans le milieu de la recherche n'aurait pu exister sans le soutien financier de la région Rhône Alpes, et plus précisément le dispositif des ARC de recherche que je remercie. Je remercie aussi la direction de l'ISA, Nathalie SCHILKNECHDT et Christophe MORELL, pour m'avoir permis d'effectuer ma thèse au sein de l'institut.

Je tiens ensuite à remercier mon directeur de thèse, le Pr. Jean-Marc LANCELIN, pour m'avoir accordé sa confiance sur ce projet pluridisciplinaire alors que je n'avais même pas encore mon master en poche ! J'en profite également pour remercier chaleureusement le Dr. Florence GUILLIERE pour son encadrement : j'apprends vite, mais il faut m'expliquer longtemps ! Merci Florence pour ton aide sur la cinétique enzymatique, la RMN, la rédaction et l'enseignement ! Merci d'avoir été là et de m'avoir soutenue dans les moments difficiles, je ne sais pas si je serai arrivée au bout de tout ça sans ton aide. J'ai appris beaucoup de choses à vos côtés à tous les deux, et je vous en remercie.

Ce travail a pu être mené à bien grâce à plusieurs collaborations. Ainsi je souhaite remercier sincèrement le Dr. Vittorio LIMONGELLI pour l'aide qu'il nous a apporté concernant la funnel-métadynamique, ainsi que le Pr. Bernard KNOOPS pour nous avoir fourni les construits qui nous ont permis de produire nos protéines d'intérêt. Je remercie plus particulièrement le Dr. Sébastien VIDAL pour sa disponibilité, et la confiance dont il a fait preuve en m'autorisant à venir effectuer des synthèses dans son laboratoire. Ça aurait pu mal tourner lorsque l'on connaît mon passé de chimiste organicienne, mais tu as su me guider juste ce qu'il faut pour que je prenne mes marques et ça a marché !

Je remercie bien sûr les personnes qui ont fait partie de l'équipe INTERACT au cours de ma thèse : Corinne, Melissa, Olivier et Maggy, pour les discussions que nous avons pu avoir à la cantine ou au laboratoire. Marie je te remercie pour ton aide précieuse au laboratoire (les journées cinétique particulièrement), pour tes ragots, et tes dessins ! Je n'oublie pas tous les stagiaires qui ont contribué aux résultats présentés dans ce manuscrit : Sabrina, Nicolas, Alexandre(s), Saloua, ... Merci pour votre bonne humeur, pour votre application dans ce projet, et pour vos gâteaux ! Je vous souhaite à tous un avenir brillant et épanouissant. Je remercie également les personnes rencontrées à l'institut pour les moments de convivialité partagés: les doctorants et post-doc d'ISA-ToPE, Sophie, Maria, ainsi que Eric, Fred, Maï et David pour tout le travail que vous fournissez au quotidien, et votre bonne humeur.

J'en arrive aux remerciements personnels, à commencer par mes amis de longue date, le groupe d'irréductibles (Marine, Ylan, Débo, Marie, Camille, Gaëlle, Aurélie, Clément, Léa et Benji), qui malgré les années (déjà 10 ans...voire plus pour certain(e)s) et la distance sont toujours présents de près ou

de loin. Je ne peux pas ne pas remercier les copains d'IUT (Yann, Flo, Laura, Quentin, Camille, Marine,...) qui ont été à la base de mon parcours étudiant en chimie et sont toujours là 6 ans après pour les moments importants ; et d'Orléans (Mehdi, Justine, Mathieu, Flavien, Loulou, Romain(s), Antho, Fiona, Pichon, Marion, Johanna, Cyriel, Seb, Wilou, Morgan, Vivi, Thomas, Lulas, et ceux que j'oublie) pour toutes les soirées passées ensemble, le BDE, la Fédéo, et pour votre amitié. J'en arrive aux lyonnais : Vinh, Yoann, Val, Elena, Aaron, Thiago, Nico : merci pour les apéros Toï-Toï après le boulot et les soirées en ville ou à la maison. Je souhaite aussi remercier mon petit Baudoin : sans cette thèse je ne t'aurais jamais rencontré... imagine tout ce qu'on aurait raté ! Merci pour tous ces moments partagés ensemble, et tous ces délires au quotidien qui ne font que commencer. Pour finir, et non des moindres, merci à celle qui m'a soutenue, subie, encouragée, et réveillée le matin quand j'étais en retard. Merci Lucie d'avoir partagé ces 3 ans avec moi. Merci pour les soirées TV, les concerts, les week-ends, les vacances, les soirées tout court, et j'en passe.... merci pour tout !

Enfin, rien de tout cela n'aurait été possible sans le soutien de ma famille. A mes parents : merci de m'avoir toujours soutenue dans mes choix, dans mes échecs et mes réussites, et merci de m'avoir permis d'arriver jusqu'ici aujourd'hui. Merci aussi au reste de la famille pour votre soutien, votre confiance et votre présence.

J'espère vous rendre fiers.

TABLE DES MATIERES

INTRODUCTION GENERALE ET PRESENTATION DES OBJECTIFS	1
CHAPITRE I.....	5
INTRODUCTION	5
I. ETUDE D'EQUILIBRES D'INTERACTIONS PROTEINE-LIGAND REVERSIBLES	5
<i>I. 1. Les caractéristiques générales d'un équilibre d'interaction réversible.....</i>	<i>5</i>
I. 1. 1 L'équilibre thermodynamique.....	5
I. 1. 2 Notion d'énergie libre standard d'interaction	8
I. 1. 3 Cinétique de l'interaction.....	8
I. 1. 4 La théorie de l'état de transition.....	11
I. 1. 5 Les différentes interactions mises en jeu.....	13
<i>I. 2. Méthodes physico-chimiques pour l'étude des équilibres d'interactions</i>	<i>15</i>
I. 2. 1 Principales méthodes d'étude de l'interaction protéine-ligand.....	15
I. 2. 2 Utilisation de la résonance magnétique nucléaire (RMN) en solution	18
I. 2. 3 Observation des signaux RMN provenant de la protéine : l'expérience HSQC.....	20
I. 2. 4 Observation des signaux RMN provenant du ligand : l'expérience STD.....	20
<i>I. 3. Modélisation utilisant la dynamique moléculaire</i>	<i>23</i>
I. 3. 1 Les bases de la modélisation moléculaire.....	23
I. 3. 2 La dynamique moléculaire	26
I. 3. 3 Prédiction de l'affinité protéine-ligand par MD.....	27
I. 3. 4 De la dynamique moléculaire à la funnel-métadynamique.....	29
II. COMMENT DECOUVRE-T-ON DES MEDICAMENTS ?	37
<i>II. 1. Conception de molécules bioactives</i>	<i>37</i>
II. 1. 1 Les différentes approches existantes.....	37
II. 1. 2 Les étapes du processus de conception.....	38
<i>II. 2. Identification et optimisation de molécules affines (hits).....</i>	<i>39</i>
II. 2. 1 Le criblage haut débit.....	39
II. 2. 2 La méthode des fragments (FBDD)	39
II. 2. 3 Les différentes méthodes analytiques pour le criblage de molécules.....	40
II. 2. 4 Optimisation des molécules hits	41
<i>II. 3. La dynamique moléculaire pour la conception de molécules affines</i>	<i>43</i>
II. 3. 1 Le criblage virtuel.....	44
II. 3. 2 Nouvelles méthodes basées sur la dynamique moléculaire.....	46
III. MODELE DE L'ETUDE : LA PEROXYREDOXINE5 HUMAINE.....	49
<i>III. 1. Les peroxyrédoxines</i>	<i>49</i>
III. 1. 1 Classification des peroxyrédoxines	49
III. 1. 2 Activité des peroxyrédoxines	55
III. 1. 3 Modèle : la peroxyrédoxine 5 humaine	59

III. 2.	<i>Une implication biologique récemment proposée : l'ischémie</i>	61
III. 2. 1	Les attaques vasculaires cérébrales	61
III. 2. 2	Le rôle des peroxyrédoxines dans l'inflammation post-ischémique	64
III. 3.	<i>Inhibition des peroxyrédoxines</i>	67
III. 3. 1	Cinétique enzymatique	67
III. 3. 2	Mesure de l'activité enzymatique	70
III. 3. 3	Le mécanisme d'inhibition enzymatique	70
III. 3. 4	La réaction enzymatique des hPRDXs	73
III. 3. 5	Inhibiteurs connus de peroxyrédoxines humaines	74
CHAPITRE II		77
MATERIEL ET METHODES		77
I.	PROTOCOLES DE MODELISATION	77
I. 1.	<i>Modélisation d'un système protéine-ligand</i>	77
I. 2.	<i>Simulation de dynamique non biaisée</i>	78
I. 3.	<i>Simulation de FM</i>	78
I. 4.	<i>Modélisation d'un complexe PL en présence de substrat (H₂O₂)</i>	79
I. 5.	<i>Estimation de la sélectivité</i>	79
I. 6.	<i>Clustering géométrique</i>	79
I. 7.	<i>Docking</i>	80
II.	PROTOCOLES DE PRODUCTION ET DE PURIFICATION DES PROTEINES	81
II. 1.	<i>Expression des isoformes 1, 2 et 5 des peroxyrédoxines humaines recombinantes</i>	81
II. 2.	<i>Expression de protéines marquées</i>	81
II. 3.	<i>Purification des isoformes 1, 2 et 5 des peroxyrédoxines humaines</i>	82
II. 4.	<i>Electrophorèse sur gel SDS</i>	83
II. 5.	<i>Expression et purification de la thioredoxine (Trx) et de la thioredoxine réductase (TrxR)</i>	86
III.	SYNTHESES D'INHIBITEURS	87
III. 1.	<i>Synthèse du 4-néopentylbenzene-1,2-diol</i>	87
III. 1. 1	1-(3,4-diméthoxyphényl)-2,2-diméthylpropan-1-one [30314-46-6]	88
III. 1. 2	4-(2,2-diméthylpropyl)-1,2-diméthoxybenzène [3]	88
III. 1. 3	4-(2,2-diméthylpropyl)benzène-1,2-diol [4]	89
III. 2.	<i>Synthèse du 4-néopentyl-5-phénylaminobenzène-1,2-diol</i>	89
III. 2. 1	4-(2,2-diméthylpropyl)-1,2-diméthoxybenzène [3]	90
III. 2. 2	1-bromo-2-(2,2-diméthylpropyl)-4,5-diméthoxybenzène [4]	90
III. 2. 3	2-(2,2-diméthylpropyl)-4,5-diméthoxy-N-phénylaniline [5]	91
III. 2. 4	4-(2,2-diméthylpropyl)-5-(phénylamino)benzène-1,2-diol [6]	91
IV.	SPECTROSCOPIE	92
IV. 1.	<i>Caractérisation de molécules organiques par RMN et SM</i>	92
IV. 2.	<i>Expériences 2D ¹⁵N HSQC</i>	92
IV. 3.	<i>Expériences STD</i>	94
IV. 4.	<i>Mesure de l'affinité protéine-ligand par RMN</i>	94

V.	EXPERIENCES DE CINETIQUE ENZYMATIQUE	96
V. 1.	Détermination expérimentale de K_M et V_{max}	96
V. 2.	Evaluation de l' IC_{50}	96
V. 3.	Etude du mécanisme d'inhibition	97
CHAPITRE III	99
RESULTATS ET DISCUSSION	99
I.	LA FUNNEL-METADYNAMIQUE ET LA RMN EN SOLUTION POUR L'ESTIMATION DE L'AFFINITE PROTEINE-LIGAND.....	101
I. 1.	Le protocole de FM pour l'étude du système hPRDX5	101
I. 1. 1	Préparation du système	101
I. 1. 1	Processus d'équilibration	104
I. 1. 2	Lancement d'une dynamique non biaisée.....	105
I. 1. 3	Mise en place d'un calcul de FM.....	105
I. 2.	Détermination de l'affinité protéine-ligand par FM	108
I. 2. 1	Dynamique non biaisée NVT.....	108
I. 2. 2	Funnel-métadynamique {FM}.....	111
I. 2. 3	Vue détaillée de l'interaction protéine-ligand par FM.....	115
I. 3.	Corrélation de l'affinité calculée par FM et déterminée par RMN	117
I. 3. 1	Détermination de l'affinité protéine-ligand par RMN	117
I. 3. 2	Etude des modes de liaison.....	121
I. 3. 3	Origines de la différence d'affinité entre CAT et 4MEC.....	123
I. 3. 4	Comparaison des affinités calculées et expérimentales	125
II.	OPTIMISATION DE LIGANDS GUIDEE PAR DYNAMIQUE MOLECULAIRE	127
II. 1.	Variations structurales et homologues du catéchol	127
II. 1. 1	Influence des hydroxyles et de leur disposition relative sur le ligand	129
II. 1. 2	Influence de l'aromaticité	133
II. 1. 3	Etude du substituant en position 4.....	135
II. 1. 4	Modification de la chaîne carbonée aliphatique	138
II. 1. 5	Etude de l'affinité du dérivé 4- <i>tert</i> -butylcatéchol	142
II. 2.	Conception et étude du 4- <i>néo</i> -pentylcatéchol.....	147
II. 2. 1	Conception du 4- <i>néo</i> -pentylcatéchol.....	147
II. 2. 2	Etude dynamique et statistique de l'interaction 4NPC-hPRDX5.....	149
II. 2. 3	Synthèse du 4- <i>néo</i> -pentylcatéchol.....	153
II. 2. 4	Etude expérimentale par RMN.....	153
II. 3.	Optimisation du 4- <i>néo</i> -pentylcatéchol par MD	155
II. 3. 1	Etude du 6-(<i>néo</i> pentyl)biphényl-3,4-diol et de dérivés.....	156
II. 3. 2	Etude de dérivés du 4- <i>néo</i> pentyl-5-(phénylamino)catéchol par FM.....	162
II. 4.	La FM pour l'optimisation des petites molécules bioactives	167
II. 4. 1	Comparaison avec le docking.....	170
II. 4. 2	Spécificité conformationnelle des différents ligands.....	172
III.	ETUDE DE L'INHIBITION DE HPRDX5 PAR CINETIQUE ENZYMATIQUE	175

III. 1. 1	Etude de l'interaction de hPRDX5 avec son substrat.....	176
III. 1. 2	Etude de l'inhibition de hPRDX5 par les dérivés catéchols.....	178
III. 1. 3	Estimation du mécanisme d'inhibition par FM.....	183
III. 1. 4	Etude de la sélectivité des dérivés catéchols.....	186
CHAPITRE IV	195
CONCLUSION ET PERSPECTIVES	195
I.	CONCLUSION	195
II.	PERSPECTIVES	197
ANNEXES	217
PUBLICATIONS	231

TABLE DES FIGURES

FIGURE 1. ISOTHERME DE LIAISON D'UN COMPLEXE PROTEINE-LIGAND.....	7
FIGURE 2 : PROFIL ENERGETIQUE D'UNE REACTION DE LIAISON D'UN COMPLEXE PL.....	9
FIGURE 3. SCHEMA REPRESENTATIF DU MODELE « CLE-SERRURE ».....	13
FIGURE 4. SCHEMA REPRESENTATIF DU MODELE DE « L'AJUSTEMENT INDUIT ».....	14
FIGURE 5. SCHEMA REPRESENTATIF DU MODELE DE LA SELECTION CONFORMATIONNELLE.....	14
FIGURE 6. SCHEMA DU PRINCIPE DE DETECTION PAR LA TECHNIQUE SPR.....	17
FIGURE 7. SCHEMA DU PRINCIPE DE DETECTION PAR LA TECHNIQUE ITC.....	17
FIGURE 8. INFLUENCE DE LA VITESSE DE L'ECHANGE SUR LES SIGNAUX RMN.....	19
FIGURE 9. SPECTRE ¹⁵ N- ¹ H HSQC DE HPRDX5 200 MM A PH = 7.4.....	21
FIGURE 10. PRINCIPE DE L'EXPERIENCE RMN 1D ¹ H DE TRANSFERT DE SATURATION (STD).....	22
FIGURE 11. VUE BIDIMENSIONNELLE D'UNE CELLULE DE SIMULATION REPLIEE DANS LES TROIS DIMENSIONS DE L'ESPACE.....	24
FIGURE 12. VISUALISATION DES DIFFERENTES COMPOSANTES DES TERMES ENERGETIQUES PRISES EN COMPTE DANS LE POTENTIEL D'UNE MOLECULE EN MECANIQUE MOLECULAIRE POUR LA DEFINITION DES CHAMPS DE FORCE.....	25
FIGURE 13. REPRESENTATION GRAPHIQUE DES DIFFERENTS MODELES DE SOLVANT EXPLICITE DEFINIS SELON LEUR GEOMETRIE ET LE NOMBRE DE SITES UTILISES.....	30
FIGURE 14. ILLUSTRATION SCHEMATIQUE DES PROGRES DES PROCESEURS GRAPHIQUES A TRAVERS L'ESSOR DE LA SOCIETE NVIDIA®.	31
FIGURE 15. PRINCIPE DE LA METADYNAMIQUE.....	32
FIGURE 16. REPRESENTATION SCHEMATIQUE DE LA DYNAMIQUE MOLECULAIRE PAR <i>UMBRELLA SAMPLING</i>	34
FIGURE 17. REPRESENTATION SCHEMATIQUE DU POTENTIEL EXTERNE FIXE UTILISE EN FM, ADAPTE AU PHENOMENE DE LIAISON PROTEINE-LIGAND.....	35
FIGURE 18. SCHEMA DES ETAPES PRINCIPALES DU PROCEDE DE CONCEPTION D'UNE MOLECULE BIOACTIVE.....	38
FIGURE 19. STRATEGIE D'EVOLUTION DE FRAGMENT POUR LA CONCEPTION D'UNE MOLECULE <i>DRUG-LIKE</i>	42
FIGURE 20. SCHEMA DE LA STRATEGIE DE LIAISON DE FRAGMENTS PERMETTANT LA CONCEPTION D'UNE MOLECULE <i>DRUG-LIKE</i>	43
FIGURE 21. REPRESENTATION SCHEMATIQUE DES DIFFERENTES APPROCHES NUMERIQUES POUR LA CONCEPTION DE MOLECULES BIOACTIVES.....	45
FIGURE 22. ALIGNEMENT DE SEQUENCE DES 6 ISOFORMES DE HPRDX ET VISUALISATION DES ACIDES AMINES CONSERVES.....	51
FIGURE 23. REPRESENTATIONS DES RESIDUS CONSERVES DANS LES HPRDX SUR LE MODELE STRUCTURAL DE HPRDX5 ET VISUALISATION DE L'ACCESSIBILITE DE LA Cp.....	52
FIGURE 24. STRUCTURES DES HPRDX ET DE LEURS INTERFACES DANS LES STRUCTURES QUATERNAIRES.....	53
FIGURE 25. MECANISME CATALYTIQUE DES PRDX EN FONCTION DE LEUR CLASSE.....	54
FIGURE 26. STRUCTURES OXYDEE ET REDUITE DE LA TRX HUMAINE, EN EQUILIBRE D'OXYDOREDUCTION.....	56
FIGURE 27. SCHEMA RECAPITULATIF DU CYCLE CATALYTIQUE DE REGENERATION DES HPRDX 1 A 5 PAR LA TRX ET LA TRXR AU TRAVERS DU NADPH.....	57
FIGURE 28. SCHEMA DES 5 ETAPES DU CYCLE CATALYTIQUE DE REDUCTION DES PEROXYDES PAR LES PRDX.....	58
FIGURE 29. VISUALISATION DE L'HOMODIMERE HPRDX5 ET DE SON SITE ACTIF.....	60

FIGURE 30. REPRESENTATION SCHEMATIQUE DES DEUX TYPES D'ACCIDENTS VASCULAIRES CEREBRAUX.....	62
FIGURE 31. VISUALISATION DU MOTIF 68-90 DE HPRDX5 IMPLIQUE DANS LA PRODUCTION DE CYTOKINES INFLAMMATOIRES.	65
FIGURE 32. VISUALISATION DE LA CASCADE ISCHEMIQUE SURVENANT APRES UN AVC.	66
FIGURE 33. REPRESENTATIONS GRAPHIQUES DES N_i EN FONCTION DE LA CONCENTRATION EN SUBSTRAT POUR DES ENZYMES MICHAELIENNES, ET DETERMINATION GRAPHIQUE DE K_M ET V_{MAX}	69
FIGURE 34. SCHEMA DES DIFFERENTS MECANISMES D'INHIBITION REVERSIBLE PRESENTES.....	72
FIGURE 35. STRUCTURE DU COMPLEXE REACTIONNEL INITIAL, ET DE L'ETAT DE TRANSITION DANS LA REACTION ENZYMATIQUE DE REDUCTION DE H_2O_2 CATALYSEE PAR HPRDX.	74
FIGURE 36. STRUCTURES DU NOYAU <i>O</i> -PYROCATECHOL ET DE MOLECULES NATURELLES POSSEDANT SON MOTIF.	76
FIGURE 37. STRUCTURE DE L'HOMODIMERE DE HPRDX2 UTILISE EN FM.	80
FIGURE 38. EXEMPLE DE GEL SDS OBTENU PAR ELECTROPHORESE EN CONDITIONS DENATURANTES POUR LA PRODUCTION DE HPRDX5 NON MARQUEE.	84
FIGURE 39. SCHEMA REACTIONNEL D'OBTENTION DU 4NPC [4] A PARTIR DU 1,2-DIMETHOXYBENZENE [1].	88
FIGURE 40. SCHEMA REACTIONNEL D'OBTENTION DES LIGANDS [6] ET [6'] A PARTIR DE [2].	90
FIGURE 41. VISUALISATION DU SYSTEME HPRDX5-CATECHOL UTILISE EN MD ET FM, APRES LES ETAPES DE SOLVATATION ET D'AJOUT DES SELS.	103
FIGURE 42. GRAPHIQUE REPRESENTANT L'EVOLUTION DES DIFFERENTS PARAMETRES INDIQUES DU SYSTEME EN FONCTION DU TEMPS, DURANT L'ETAPE D'EQUILIBRATION D'UN SYSTEME HPRDX5 EN SOLVANT EXPLICITE (TIP3P) ETUDIE PAR MD.	104
FIGURE 43. PARAMETRAGE DU POTENTIEL DE CONTRAINTE SPATIALE DU LIGAND UTILISE EN FM.	106
FIGURE 44. DEFINITION DE LA VARIABLE COLLECTIVE CV3 SUIVIE EN FM.	107
FIGURE 45. DYNAMIQUE NVT DE HPRDX5 COMPLEXEE AU DTT OXYDE EN PRESENCE DE 6 CATECHOLS.	109
FIGURE 46. LES DEUX MODES DE LIAISON DU CATECHOL OBSERVES EN DYNAMIQUE NVT AVEC HPRDX5.	110
FIGURE 47. ANALYSE DE LA TRAJECTOIRE NVT DE HPRDX5 COMPLEXE AU DTT OXYDE EN PRESENCE DE 6 LIGANDS 4- METHYL CATECHOL AU COURS DE LA DYNAMIQUE NVT.	111
FIGURE 48. EVOLUTION DES DIFFERENTES VARIABLES COLLECTIVES AU COURS DU TEMPS DE SIMULATION FM DES SYSTEMES CAT- HPRDX5 (HAUT) ET 4MEC-HPRDX5 (BAS).	112
FIGURE 49. COURBES $W(Z)=F(CV1)$ ET SURFACES $W(Z)=F(CV1,CV2)$ DES COMPLEXES HPRDX5-CAT ET HPRDX5-4MEC ETUDIES PAR FM.	113
FIGURE 50. DEPOT DU POTENTIEL DE BIAIS HILLS AU COURS DU TEMPS DE SIMULATION.	114
FIGURE 51. EVOLUTION DE L'ENERGIE LIBRE DE LIAISON (KCAL.MOL ⁻¹) AU COURS DES 500 NS DE SIMULATION FM POUR LES SYSTEMES CAT-HPRDX5 ET 4MEC-HPRDX5.	115
FIGURE 52. ETUDE DETAILLEE DE L'INTERACTION DE CAT ET 4MEC AVEC HPRDX5 PAR FM.	116
FIGURE 53. CLUSTERING GEOMETRIQUE DES COMPLEXES HPRDX5-4MEC OBTENUS PAR FM.	116
FIGURE 54. INFLUENCE DU DMSO SUR LES SIGNAUX RMN DE ¹⁵ N-HPRDX5 A 28°C, DANS DU PBS PH=7.4, ENREGISTRE A 600 MHZ.	119
FIGURE 55. SPECTRE ¹ H- ¹⁵ N HSQC DE HPRDX5 (400 MM, 600 MHZ, 28°C, PH=7.4) EN PRESENCE D'AJOUTS CROISSANTS DE CAT. VISUALISATION DES CSPS INDUITES PAR LE LIGAND.	120
FIGURE 56. COURBES ISOTHERMES REPRESENTANT LA QUANTITE DE CSP INDUITE PAR CAT ET 4 MEC SUR LES SIGNAUX ¹ H- ¹⁵ N DE HPRDX5 A 28°C ET PH=7.4 DANS UN TAMPON PBS.	121

FIGURE 57. CARTOGRAPHIES D'ÉPITOPES OBTENUES PAR RMN STD POUR LES LIGANDS CAT ET 4MEC EN INTERACTION AVEC HPRDX5, A 28°C DANS DU TAMPON PBS, ENREGISTRE A 600 MHZ.	122
FIGURE 58. HISTOGRAMME CUMULE REPRESENTANT LES CSPS INDUITES PAR CAT ET 4MEC SUR LES SIGNAUX RMN DE HPRDX5, POUR UN RAPPORT $[L]/[P]$ CORRESPONDANT A $2K_D$	124
FIGURE 59. STRUCTURES DES MOLECULES FRAGMENTS DONT L'AFFINITE POUR HPRDX5 A PU ETRE DETERMINEE PAR FM, POUR L'ETUDE DE L'INTERACTION PL.	127
FIGURE 60. CINQ MODES DE LIAISON POSSIBLES DU LIGAND GALLATE D'ETHYLE AVEC LE THIOLATE CYS47.	131
FIGURE 61. ANALYSE DES MODES DE LIAISON DU LIGAND GALLATE D'ETHYLE AVEC HPRDX5 PAR FM.	132
FIGURE 62. HISTOGRAMME REPRESENTANT LES CSPS INDUITES PAR LE LIGAND GAL (7MM) SUR LES SIGNAUX DE ^{15}N -HPRDX5 (200 MM), A 600 MHZ, 28°C, PH=7.4.	133
FIGURE 63. STRUCTURE DES ANALOGUES <i>CIS</i> ET <i>TRANS</i> DU CYCLOHEXANE-1,2-DIOL ET DU DTT SOUS SA FORME OXYDEE CYCLIQUE.	134
FIGURE 64. HISTOGRAMMES DES CSPS INDUITES PAR L'INTERACTION DES ISOMERES <i>CIS</i> ET <i>TRANS</i> DU CYCLOHEXANE-1,2-DIOL SUR ^{15}N -HPRDX5, ANALYSES PAR RMN A 600 MHZ, 28°, PH=7.4.	135
FIGURE 65. CLUSTERS DES CONFORMATIONS DE LIAISON OBSERVES POUR L'ACIDE (3,4-DIHYDROXYPHENYL) ACETIQUE AVEC HPRDX5 PAR FM.	137
FIGURE 66. VISUALISATION DE L'ENERGIE LIBRE DE LIAISON DU COMPLEXE 4-(METHYLAMINO)CATECHOL-HPRDX5 AU COURS DE LA SIMULATION FM.	138
FIGURE 67. ANALYSE FM DU COMPLEXE HPRDX5-4ECAT : FES CV1-CV3 ET ΔG_B°	139
FIGURE 68. DIFFERENCES EN TERMES D'ENERGIE DE LIAISON, ENTRE LES COMPLEXES FORMES PAR LES LIGANDS 4- <i>ISO</i> -PROPYLCATECHOL ET 4- <i>ISO</i> -BUTYLCATECHOL AVEC HPRDX5.	141
FIGURE 69. ANALYSE DE LA SIMULATION DU COMPLEXE 4TBC AVEC HPRDX5 PAR FM.	143
FIGURE 70. ANALYSE DES 3 MODES DE LIAISON DU LIGAND 4TBC AVEC HPRDX5 PAR FM.	144
FIGURE 71. INTERACTION DU 4TBC AVEC ^{15}N -HPRDX5 (200 UM) ETUDIEE PAR RMN ^{15}N -HSQC, A 600 MHZ, 28°C, PH=7.4.	145
FIGURE 72. IMPLICATION DES DIFFERENTS PROTONS AROMATIQUES DU 4TBC DANS SON INTERACTION AVEC HPRDX5 (10 UM) PAR EXPERIENCE STD, A 600 MHZ, 28°C, PH=7.4.	146
FIGURE 73. REPRESENTATION SCHEMATIQUE DU PRINCIPE DE LIAISON HYDROGENE ECRANTEE DU SOLVANT.	148
FIGURE 74. ANALYSE DE LA SIMULATION DE L'INTERACTION DU 4- <i>NEO</i> -PENTYLCATECHOL AVEC HPRDX5 PAR FM.	150
FIGURE 75. ANALYSE DES 4 MODES DE LIAISON DU 4- <i>NEO</i> -PENTYLCATECHOL AVEC HPRDX5 PAR FM.	150
FIGURE 76. DYNAMIQUE INTERNE DU SITE ACTIF PAR SUIVI DE DISTANCES INTERATOMIQUES EN FONCTION DU TEMPS DE SIMULATION (NS) POUR LES 4 LIGANDS SUIVANTS : CAT, 4MEC, 4TBC ET 4NPC.	152
FIGURE 77. RESULTATS DE L'AFFINITE DU 4NPC POUR HPRD5 OBTENUS EXPERIMENTALEMENT PAR RMN EN SOLUTION.	154
FIGURE 78. RESULTATS OBTENUS PAR EXPERIENCES RMN STD POUR L'INTERACTION DE 4NPC AVEC HPRDX5.	155
FIGURE 79. DIFFERENTES INTERACTIONS NON COVALENTES ATTRACTIVES POSSIBLES ENTRE DEUX CYCLES AROMATIQUES : LE « <i>π-STACKING</i> ».	156
FIGURE 80. SURFACE D'ENERGIE LIBRE DU COMPLEXE 6NBD -HPRDX5 AU COURS DE LA SIMULATION FM.	158
FIGURE 81. ANALYSE DE L'INTERACTION DU 4-NEOPENTYL-5-(PHENYLAMINO)CATECHOL AVEC HPRDX5.	160
FIGURE 82. DYNAMIQUE DU SITE ACTIF EN PRESENCE DU 4-NEOPENTYL-5-(PHENYLAMINO)CATECHOL.	160

FIGURE 83. SURFACES D'ÉNERGIE LIBRE DE LIAISON DES LIGANDS HIT1 ET HIT2 EN INTERACTION AVEC HPRDX5 LORS DE LA SIMULATION FM.....	163
FIGURE 84. ANALYSE DES MODES DE LIAISONS MAJORITAIRES DU HIT 3 DANS LE SITE CATALYTIQUE DE HPRDX5, ET VISUALISATION DE L'ÉNERGIE DE LIAISON DU COMPLEXE ASSOCIE.	166
FIGURE 85. STRUCTURES DES LIGANDS DONT L'AFFINITE POUR HPRDX5 A PU ETRE ESTIMEE PAR FM LORS DE CETTE ETUDE.	169
FIGURE 86. HISTOGRAMME EMPILE REPRESENTANT LA CSP TOTALE SUBIE PAR HPRDX5 SOUS L'INFLUENCE DES LIGANDS CATECHOLS CAT, 4MEC, 4TBC ET 4 NPC POUR DES CONCENTRATIONS EN LIGAND ET PROTEINE EQUIVALENTES A 67% DE LA SATURATION TOTALE.....	174
FIGURE 87. MECANISME GENERAL DE L'INHIBITION REVERSIBLE DES HPRDX.....	175
FIGURE 88. SIMULATION DU PEROXYDE D'HYDROGENE AVEC HPRDX5 : ENERGIE DU COMPLEXE APPARENT ES.....	177
FIGURE 89. ETUDE DE L'INHIBITION DU 4NPC POUR HPRDX5.....	179
FIGURE 90. ANALYSE DU MECANISME D'INHIBITION DE HPRDX5 PAR LE 4NPC.	180
FIGURE 91. COMPARAISON DES IC ₅₀ OBTENUS POUR LES DIFFERENTS DERIVES CATECHOLS ETUDIES AVEC HPRDX5.....	182
FIGURE 92. REPRESENTATION SCHEMATIQUE DES CONFORMATIONS DE LIAISON (HAUT) ET FES REPRESENTANT L'ÉNERGIE LIBRE DU COMPLEXE HPRDX5-4TBC EN ABSENCE ET EN PRESENCE DE H ₂ O ₂ EN FONCTION DE CV1 ET CV2 (BAS).	184
FIGURE 93. REPRESENTATION SCHEMATIQUE DES CONFORMATIONS DE LIAISON (HAUT) ET FES REPRESENTANT L'ÉNERGIE LIBRE DU COMPLEXE HPRDX5-4NPC EN ABSENCE ET EN PRESENCE DE H ₂ O ₂ EN FONCTION DE CV1 ET CV2 (BAS).....	185
FIGURE 94. COMPARAISON DE L'INHIBITION DU LIGAND 4NPC POUR HPRDX1, HPRDX2 ET HPRDX5.....	187
FIGURE 95. ANALYSE DE L'ÉNERGIE DU COMPLEXE HPRDX2-4NPC LORS DE LA SIMULATION FM.....	191
FIGURE 96. VISUALISATION DES 4 CONFORMATIONS DE LIAISON LES PLUS FREQUEMMENT OBSERVEES AU COURS DE LA SIMULATION FM DU COMPLEXE HPRDX2-4NPC.....	192
FIGURE 97. ANALYSE DE L'ÉNERGIE DU COMPLEXE HPRDX5-PEROXYDE DE CUMENE PAR FM.	198
FIGURE 98. COURBE DE CINETIQUE EFFECTUEE AVEC LE PEROXYDE DE CUMENE COMME SUBSTRAT DE LA REACTION ENZYMATIQUE CATALYSEE PAR HPRDX5.	199

TABLE DES TABLEAUX

TABLEAU 1. COMPARAISON DES METHODES HTS ET FBDD	40
TABLEAU 2. COMPOSITION DES MILIEUX DE CULTURE.....	82
TABLEAU 3. COMPOSITION DES SOLUTIONS TAMPONS UTILISEES	83
TABLEAU 4. COMPOSITION DES GELS ET TAMPONS UTILISES EN ELECTROPHORESE	85
TABLEAU 5. EXPERIENCES RMN ET PARAMETRES D'ACQUISITION UTILISES	93
TABLEAU 6. AFFINITES DETERMINEES POUR LES LIGANDS CAT ET 4MEC POUR HPRDX5 PAR FM ET RMN.	125
TABLEAU 7. TABLEAU RECAPITULATIF DES ENERGIES DE LIAISON (ΔG°_8) ET CONSTANTES DE DISSOCIATION OBTENUES POUR LES DIFFERENTS LIGANDS ETUDIES AVEC HPRDX5, PAR FM	128
TABLEAU 8. RESULTATS D'AFFINITE DES LIGANDS 4- <i>ISO</i> -PROPYLCATECHOL ET 4- <i>ISO</i> -BUTYLCATECHOL POUR HPRDX5 ESTIMES PAR FM	140
TABLEAU 9. RESULTATS D'AFFINITE DES LIGANDS HIT1 ET HIT2 POUR HPRDX5 ESTIMES PAR FM.....	163
TABLEAU 10. BILAN DE TOUTES LES MOLECULES TESTEES PAR FM DONT L'AFFINITE POUR HPRDX5 A ETE CALCULEE	167
TABLEAU 11. RESULTATS DES ENERGIES DE LIAISON OBTENUES PAR <i>DOCKING</i> POUR LES LIGANDS CAT, 4MEC, 4TBC, 4NPC, HIT4 ET HIT4N EN COMPLEXE AVEC HPRDX5.	171
TABLEAU 12. COMPARAISON DE L'INHIBITION DE HPRDX5 PAR LES LIGANDS CATECHOLS ETUDIES.....	182
TABLEAU 13. COMPARAISON DES RESULTATS D'AFFINITE ESTIMES PAR FM POUR L'ENZYME LIBRE HPRDX5 ET POUR LE COMPLEXE ES HPRDX5-H ₂ O ₂ : CAS DES LIGANDS 4TBC ET 4NPC.	184
TABLEAU 14. AFFINITE DES DERIVES CATECHOLS POUR DIFFERENTES PROTEINES HPRDXS.....	188
TABLEAU 15. IC ₅₀ DES DERIVES 4TBC ET 4NPC POUR LES ISOFORMES HPRDX1 HPRDX2 ET HPRDX3.	189
TABLEAU 16. EVOLUTION DU T _R MOYEN MESURE PAR FM EN FONCTION DE L'AFFINITE ESTIMEE POUR HPRDX5, POUR LES PRINCIPAUX DERIVES CATECHOLS ETUDIES.....	200

LISTE DES ABREVIATIONS ET SYMBOLES

Abréviations

4B5NPC	4-benzyl-5- <i>néo</i> -pentylcatéchol
4MEC	4-méthylcatéchol
4NPC	4- <i>néo</i> -pentylcatéchol
4N5PC	4- <i>néo</i> -pentyl-5-(phenylamino)catéchol
4TBC	4- <i>tert</i> -butylcatéchol
6NBD	6-(<i>néo</i> -pentyl)biphényl-3,4-diol
AMBER	<i>Assisted Model Building with Energy Refinement</i>
Ar	Argon
ATP	Adenosine TriPhosphate
AVC	Accident Vasculaire Cérébral
[BMIM][PF ₆]	1-butyl-3-méthylimidazoliumhexafluorophosphate
CAT	Catéchol
CSP	Perturbation de déplacement chimique (<i>Chemical Shift Perturbation</i>)
CYH	Cyclohexane-1,2-diol
DMSO	Diméthylsulfoxyde
DTT	Dithiothréitol
EDTA	Ethylène Diamine Tétracétique
FF	<i>Fully Folded</i>
FM	<i>Funnel Metadynamics</i>
GAL	Gallate d'éthyle
GPU	Processeur graphique (<i>Graphic Processor Unit</i>)
hPRDX	Peroxyrédoxines humaines
hPRDX5	Peroxyrédoxine 5 humaine (<i>Human Peroxiredoxin 5</i>)
HSQC	<i>Heteronuclear Single Quantum Coherence spectroscopy</i>
IC ₅₀	Concentration inhibitrice médiane
LE	Efficacité ligand (<i>Ligand Efficiency</i>)
LU	<i>Locally Unfolded</i>
MD	Dynamique Moléculaire (<i>Molecular Dynamics</i>)
NADPH	Nicotinamide Adénine Dinucléotide Phosphate
NBS	<i>N</i> -bromosuccinimide
PA	Activateur de plasminogène
PBS	Tampon Phosphate Salin (<i>Phosphate Buffer Saline</i>)
PDB	<i>Protein Data Bank</i>
PL	Complexe protéine-ligand
PRDX	Peroxyrédoxines
PSI-BLAST	<i>Position-Specific Iterated – Basic Local Alignment Search Tool</i>
ROS / RNS	Espèces réactives oxygénées / Espèces réactives azotées (<i>Reactive Oxygen Species / Reactive Nitrogen Species</i>)
RMN	Résonance Magnétique Nucléaire

SDS	Dodécylsulfate de sodium (<i>Sodium Dodecyl Sulfate</i>)
STD	Différence de transfert de saturation (<i>Saturation Transfer Difference</i>)
TCEP	Tris(2-carboxyéthyl)phosphine
TEMED	N,N,N',N'-tétraméthyléthylènediamine
Trx/TrxR	Thiorédoxine et Thiorédoxine réductase
vdW	van der Waals
WaterLOGSY	<i>Water Ligand Observed via Gradient Spectroscopy</i>

Symboles

C_p	Cystéine peroxydatique
C_r	Cystéine de régénération
<i>C-term</i>	Extrémité C _{terminale} d'une chaîne protéique
f_{STD}	Facteur d'amplification STD
I_{1D}	Intensité d'un signal dans l'expérience 1D proton
I_{STD}	Intensité d'un signal dans l'expérience STD
k_b	Constante de Boltzmann
k_{cat} et k_{cat}'	Constantes de vitesse catalytique
K_A	Constante d'équilibre d'association
K_D	Constante d'équilibre de dissociation
K_{eq}	Constante d'équilibre
K_i et K_i'	Constantes d'inhibition d'un inhibiteur pour une enzyme
K_M	Constante de Michaelis
k_{on} et k_{off}	Constantes de vitesse d'association et de dissociation d'un complexe protéine-ligand
M, mM, μ M, nM	mol.L ⁻¹ ; mmol.L ⁻¹ ; μ mol.L ⁻¹ ; nmol.L ⁻¹
<i>N-term</i>	Extrémité N _{terminale} d'une chaîne protéique
V_0	Vitesse initiale d'une réaction enzymatique
V_{max}	Vitesse maximale d'une réaction enzymatique

LISTE DES ACIDES AMINES ET LEUR DENOMINATION

Acide aspartique	Asp	D
Acide glutamique	Glu	E
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Cystéine	Cys	C
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Méthionine	Met	M
Phénylalanine	Phe	F
Proline	Pro	P
Sérine	Ser	S
Thréonine	Thr	T
Tryptophane	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

INTRODUCTION GENERALE ET PRESENTATION DES OBJECTIFS

L'étude présentée dans ce manuscrit porte sur la compréhension de l'interaction protéine-ligand (PL) et le calcul d'affinité par dynamique moléculaire pour la conception de nouvelles molécules inhibitrices, appliqué au modèle biologique de la peroxyrédoxine 5 humaine (hPRDX5).

Cette protéine est impliquée dans diverses pathologies, et a récemment été montrée comme étant une cible thérapeutique potentielle dans l'ischémie cérébrale. Nous avons donc centré notre travail sur l'étude de cette protéine modèle dans ce contexte pharmacologique particulier. Cependant, les protocoles utilisés lors de ce projet restent applicables à de nombreux autres systèmes. Aussi, nous sommes convaincus qu'au-delà de la conception d'inhibiteurs de la peroxyrédoxine 5 humaine, ce travail aura un impact méthodologique sur l'étude et la prédiction du phénomène d'association protéine-ligand, et amènera de nouvelles perspectives dans le domaine de la chimie médicinale.

Ce projet de thèse a été réalisé en trois temps.

Tout d'abord la dynamique moléculaire (MD) a été utilisée pour la compréhension de l'interaction PL et l'étude de l'affinité de ces complexes grâce à la statistique des conformations de liaison et la cartographie possible de l'énergie en funnel-métadynamique (FM). Dans le cadre de cette première partie théorique, la corrélation des énergies libres de liaison simulées avec les résultats expérimentaux obtenus par RMN en solution est aussi réalisée.

Dans un deuxième temps, les informations tirées des études par simulation ont été utilisées pour l'amélioration de l'affinité PL. Cette partie du projet s'inscrit dans la conception et l'optimisation de molécules bioactives dont je présenterai les notions fondamentales. Depuis quelques années la conception de molécules bioactives s'appuie de plus en plus fortement sur les évolutions et les progrès de la MD afin de gagner en efficacité et en productivité pour la mise sur le marché de nouvelles molécules.

Enfin, nous avons étudié si l'affinité observée se vérifie en termes d'activité inhibitrice. Ainsi, nous avons vérifié l'inhibition de la protéine cible : la peroxyrédoxine 5 humaine (hPRDX5). Cette enzyme est impliquée dans la régulation du peroxyde d'hydrogène dans la cellule, et montrée impliquée dans l'inflammation post-ischémique cérébrale. Cette dernière partie du projet a été faite en collaboration avec M.L. Chow dans le cadre de sa thèse sur la compréhension des effets de l'inhibition des peroxyrédoxines humaines.¹

Dans le chapitre d'introduction, je présenterai les notions nécessaires à la compréhension de l'étude et des résultats obtenus au cours de ma thèse. Le chapitre d'introduction comportera trois parties distinctes. La première partie sera consacrée à l'étude d'équilibres d'interactions réversibles protéine-ligand, ainsi qu'aux différentes techniques physico-chimiques permettant la mesure des grandeurs thermodynamiques et cinétiques, et leur estimation possible par MD. La seconde partie traitera du processus de découverte de molécules bioactives par différentes approches possibles et l'utilisation récente de la MD en rapport avec elles. Enfin, la troisième et dernière partie introductive servira à présenter le modèle biologique hPRDX5 sur lequel ont été menées les expériences présentées dans ce manuscrit.

A la suite de ce chapitre introductif seront présentés les matériels et les méthodes utilisés au cours de ce projet. Les paramètres utilisés lors des simulations de MD et de FM seront détaillés, ainsi que les analyses effectuées à la suite de ces simulations. Les protocoles de production et de purification des différentes protéines utilisées ainsi que les synthèses organiques de deux inhibiteurs seront décrits. Enfin, nous décrirons les méthodes expérimentales de spectroscopie et de cinétique enzymatique utilisées.

Après avoir abordé les notions et les méthodes utiles, je présenterai les résultats obtenus au cours de ma thèse. Ce chapitre sera aussi composé de trois parties distinctes correspondant aux différents objectifs présentés précédemment. Ainsi, seront tout d'abord abordés les résultats obtenus sur l'interaction du ligand catéchol et de son dérivé 4-méthylcatéchol avec hPRDX5 par FM, ainsi que la corrélation des affinités PL obtenues lors de ces simulations avec celles déterminées expérimentalement par RMN en solution. Les résultats de l'étude de l'interaction entre d'autres dérivés du catéchol et hPRDX5 par FM seront ensuite présentés et discutés. Cette étude a mené à l'optimisation du fragment catéchol initial pour en améliorer l'affinité PL. La dernière partie de ce chapitre portera sur l'étude de l'inhibition de hPRDX5 par certains des dérivés étudiés, ainsi que les mécanismes mis en jeu lors de cette inhibition enzymatique. L'inhibition d'autres isoformes chez l'homme sera présentée afin d'estimer la possible sélectivité d'action de ces inhibiteurs. Enfin, l'utilisation de la FM comme outil pour l'estimation des mécanismes d'inhibition sera discutée.

CHAPITRE I

INTRODUCTION

I. ETUDE D'ÉQUILIBRES D'INTERACTIONS PROTÉINE-LIGAND RÉVERSIBLES

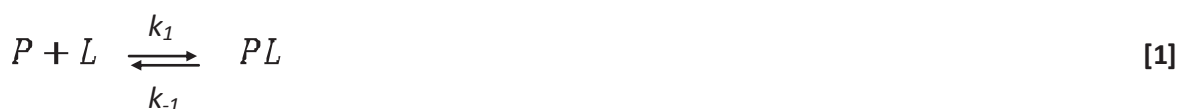
Les protéines possèdent de nombreux rôles biologiques : de la catalyse de réactions chimiques jusqu'à la régulation d'autres protéines en passant par la détection de signaux au travers de leur liaison avec d'autres entités biologiques. Le rôle central des protéines est donc d'interagir avec d'autres molécules afin de former un complexe transitoire. C'est pourquoi la compréhension et la description des interactions protéine-ligand d'un point de vue thermodynamique et cinétique est d'une importance capitale pour la compréhension de phénomènes biologiques de toutes sortes.^{2; 3} Nous définirons dans cette partie introductive les caractéristiques générales thermodynamiques et cinétiques d'un équilibre d'interaction réversible protéine-ligand, puis nous exposerons brièvement les différentes techniques macroscopiques pouvant être utilisées dans l'étude de ces équilibres réversibles, et plus particulièrement la RMN qui sera utilisée dans cette étude. Enfin, nous terminerons par présenter la modélisation de ces équilibres d'interaction en utilisant la dynamique moléculaire microscopique pour la prédiction de phénomènes dynamiques et macroscopiques.

I. 1. Les caractéristiques générales d'un équilibre d'interaction réversible

Pour étudier la complexation d'une protéine à un ligand, l'approche classique consiste à assimiler l'interaction à un équilibre chimique. Comme tout équilibre chimique, celui-ci peut se caractériser d'un point de vue thermodynamique, mais aussi d'un point de vue cinétique.

I. 1. 1 L'équilibre thermodynamique

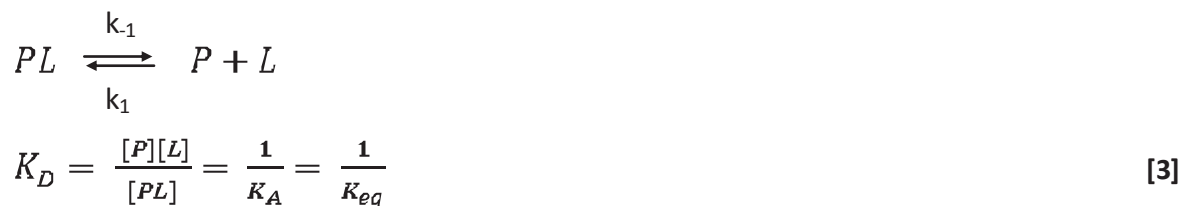
Si on se restreint dans un premier temps à considérer 1 seul ligand par protéine, l'équilibre d'interaction s'écrit :



On définit alors la constante d'équilibre K_{eq} comme le rapport des concentrations des espèces libres et liées, à l'équilibre thermodynamique, pour une température donnée.

$$K_{eq} = \frac{[PL]}{[P][L]} \quad [2]$$

Où $[PL]$ est la concentration du complexe PL, $[P]$ la concentration en protéine libre, et $[L]$ la concentration du ligand libre. Comme l'équation 1 est en direction de la formation du complexe, cette constante d'équilibre K_{eq} est aussi nommée constante d'association du complexe protéine-ligand (K_A). Cependant il est généralement plus commun de caractériser la stabilité d'un complexe PL par la valeur de sa constante de dissociation (K_D), dont la formule est simplement l'inverse de K_A , correspondant à la réaction de dissociation du complexe PL :



Ces constantes d'équilibre, bien que normalement sans unités sont le plus souvent exprimées en unités molaires (mol.L^{-1}) du fait de leur expression par rapport aux fractions libres et liées de protéine et de ligand en solution, exprimées en unités de concentration (Equation 3). Mais l'équation 3 n'est pas complète, car les concentrations devraient être exprimées par rapport aux valeurs de concentrations standard $[P]^\circ$, $[L]^\circ$ et $[PL]^\circ$ comme exprimé dans l'équation 4 suivante :

$$K_D = \frac{\left(\frac{[P]}{[P]^\circ}\right) \cdot \left(\frac{[L]}{[L]^\circ}\right)}{\left(\frac{[PL]}{[PL]^\circ}\right)} \quad [4]$$

Cependant, les concentrations standards étant égales à 1 M, elles sont communément omises, et il en résulte l'expression des constantes K_A et K_D telles qu'exprimées dans les équations 2 et 3, impliquant l'unité molaire de la concentration. C'est donc l'unité qui sera adoptée pour le reste du manuscrit. K_D correspond à la concentration en ligand pour laquelle 50% de la protéine est complexée avec le ligand. Cette valeur peut être déterminée expérimentalement à partir de l'isotherme de liaison du complexe (Figure 1) par de nombreuses techniques expérimentales que nous expliciterons par la suite.

Typiquement, lorsqu'il s'agit d'une molécule thérapeutique, il est classique d'atteindre au minimum un K_D de l'ordre du nano-molaire, soit $K_D \leq 10^{-9}$ M. En effet, une interaction de l'ordre du micro-molaire nécessiterait des doses irréalistes pouvant donner lieu à des effets secondaires indésirables.

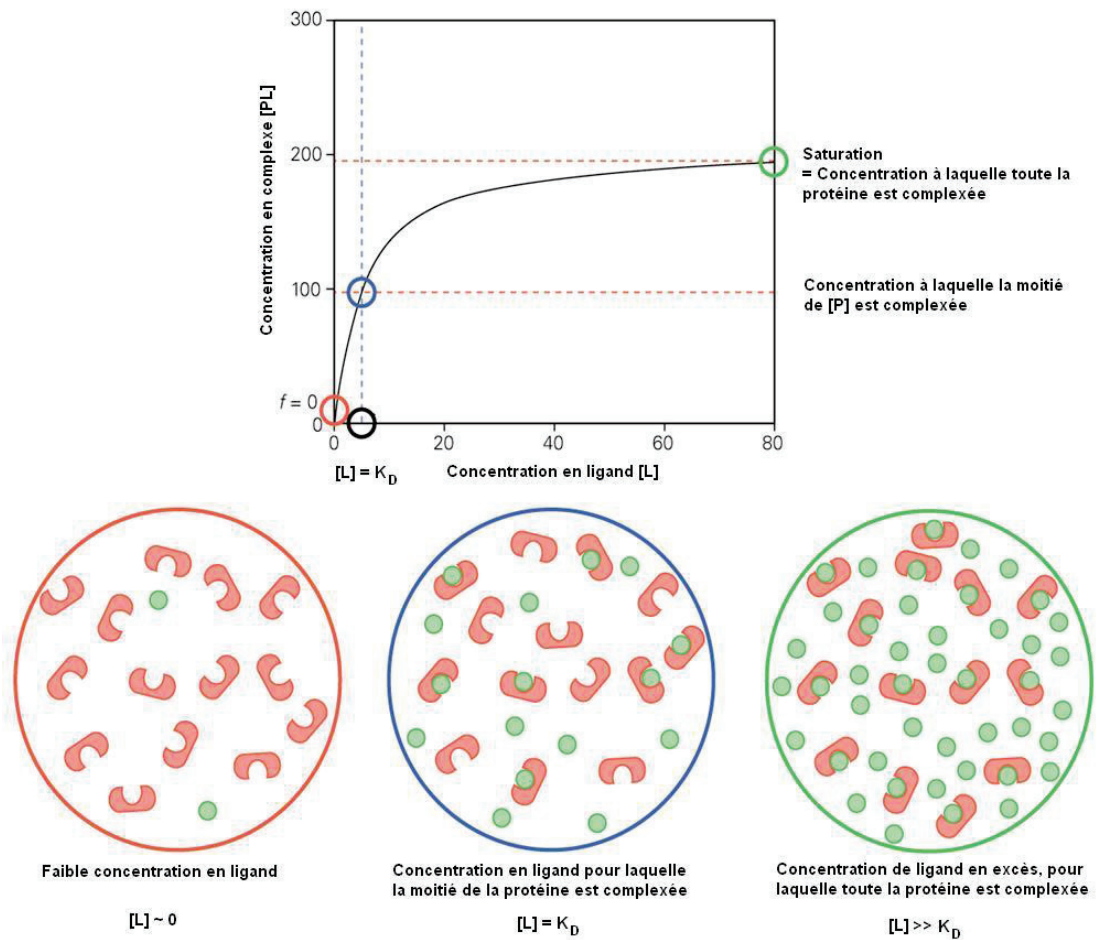


Figure 1. Isotherme de liaison d'un complexe protéine-ligand.

Ce graphique représente la valeur de la concentration en complexe $[PL]$ en fonction de la concentration en ligand $[L]$.

Figure reproduite et traduite de Kuriyan , Konforti & Wemmer. The Molecules of Life. Avec l'autorisation de Garland Publishing. 2008.⁴

Il est donc préférable d'atteindre des valeurs de K_D les plus basses possibles, afin de minimiser les effets secondaires. Cependant nous verrons par la suite que la seule considération du K_D ne permet pas de caractériser significativement l'activité biologique d'une molécule donnée.⁴

I. 1. 2 Notion d'énergie libre standard d'interaction

L'affinité d'une interaction moléculaire fait référence à la stabilité du complexe formé. Nous avons vu que l'une des grandeurs thermodynamiques qui permettait d'estimer cette affinité est la constante de dissociation du complexe protéine-ligand. Cette constante thermodynamique peut aussi être reliée à une autre grandeur : la différence d'énergie libre ΔG (appelée aussi énergie de Gibbs) entre l'état libre et l'état lié des composants de l'équilibre. Si cette différence exprimée en J (Joules) ou cal (1 cal = 4.18 J) est négative, elle démontre un processus d'association favorable qui sera spontané, avec une dissociation défavorable et donc un complexe stabilisé. Au contraire, un ΔG positif indiquera un processus défavorable dans le sens de la formation du complexe. Lorsque ΔG est exprimée dans les conditions de concentration standard (1 M), on définit alors ΔG_b° , la différence d'énergie libre absolue d'association en $\text{J}\cdot\text{mol}^{-1}$ ou $\text{cal}\cdot\text{mol}^{-1}$. ΔG_b° est lié à K_A et K_D par l'égalité suivante :

$$\Delta G_b^\circ = -RT \ln K_A = +RT \ln K_D \quad [5]$$

Avec ΔG_b° l'énergie libre standard de formation du complexe PL, R la constante des gaz parfaits ($R = 8.3 \text{ J}\cdot\text{K}^{-1}\cdot\text{mol}^{-1}$), et T la température en Kelvin.

Au vu de l'équation 5, nous pouvons entrevoir que de faibles modifications en termes d'énergie, mèneront à de grandes variations en termes de K_D , dues au terme logarithmique. Comme on peut se rendre compte d'après le diagramme de variation d'énergie libre d'une réaction chimique quelconque (Figure 2), un système évolue spontanément vers les processus favorables pour lesquels $\Delta G^\circ < 0$. Donc plus ΔG_b° va être négative, plus K_A va être grand (et K_D faible), et plus la réaction va être favorable dans le sens de la formation du complexe PL^{2;3}

I. 1. 3 Cinétique de l'interaction

Comme nous venons de le définir, une réaction pourrait être réalisée si elle est thermodynamiquement favorable. Cependant, l'équilibre thermodynamique ne procure aucune information sur la cinétique du processus. Dans le modèle à un ligand défini ci-dessus, la vitesse d'association et la vitesse de dissociation du complexe PL peuvent être décrites par deux constantes de vitesse, notées respectivement k_{on} et k_{off} .

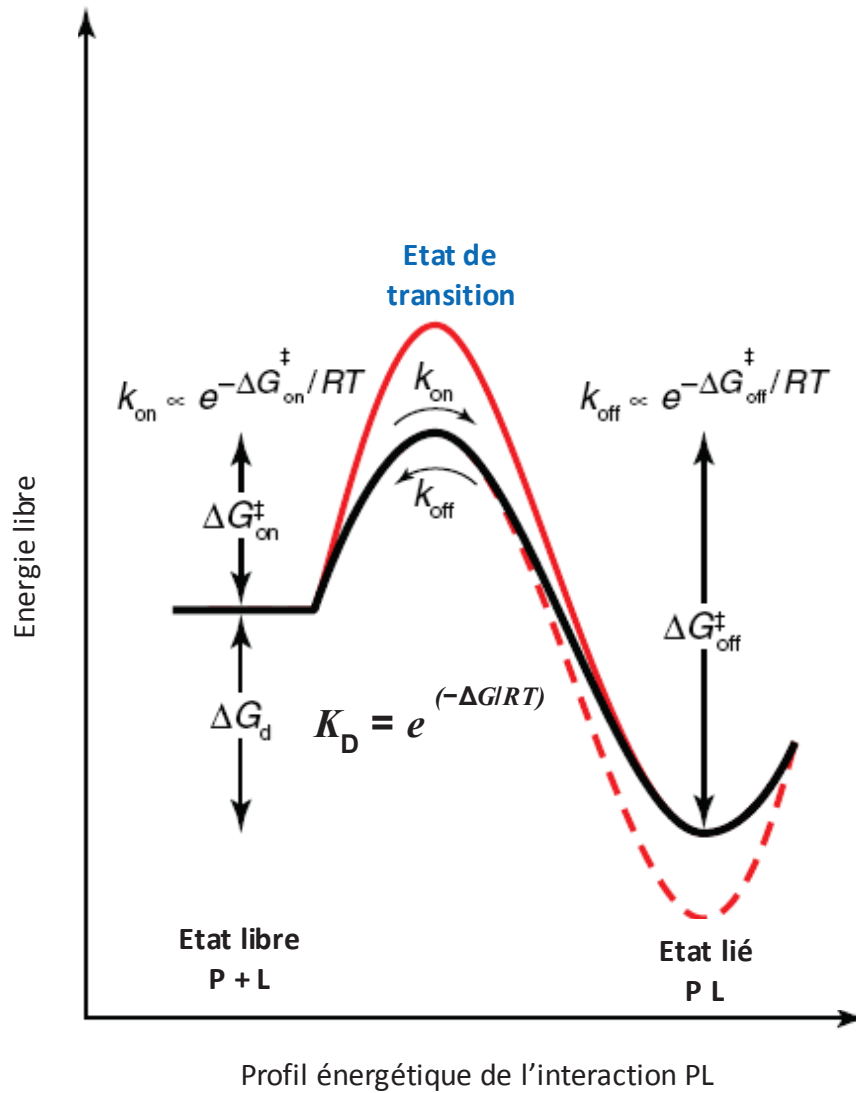


Figure 2 : Profil énergétique d'une réaction de liaison d'un complexe PL.

Profil énergétique d'un ligand (L) se liant à une protéine (P) pour former un complexe (PL). La différence d'énergie libre entre les états libres et liés ΔG , ne dépend que de ces 2 états, et détermine l'affinité du ligand pour la protéine (K_D). Les constantes de vitesse d'association et de dissociation k_{on} et k_{off} dépendent des différences d'énergie libre entre les états libres et liés, et l'état de transition, respectivement ΔG_{on}^\ddagger et ΔG_{off}^\ddagger . Avec R la constante des gaz parfaits, T la température en Kelvin.

Schéma reproduit et modifié à partir de Pan et al. 2013. Drug Discovery Today 18(13-14): 667-673 avec autorisation de Elsevier.⁵

Elles sont reliées à K_D qui est le rapport direct de k_{off}/k_{on} :

$$K_D = \frac{k_{off}}{k_{on}} \quad [6]$$

Donc,

$$\frac{k_{off}}{k_{on}} = \exp\left(\frac{\Delta G_b^\ddagger}{RT}\right) \quad [7]$$

Soit à l'équilibre chimique, en reprenant l'équation 3 et en utilisant la loi empirique d'action de masse,⁶ on obtient l'égalité suivante reliant les vitesses v de l'équilibre d'interaction aux concentrations en protéine libre, en ligand libre, et en complexe PL :

$$v = [P][L] k_{on} = [PL] k_{off} \quad [8]$$

Avec k_{on} exprimée en $M^{-1}.s^{-1}$ et k_{off} exprimée en s^{-1} .

Ces constantes de vitesse k_{on} et k_{off} reflètent la vitesse à laquelle se font l'association et la dissociation du complexe PL. k_{off} peut aussi être relié à un autre facteur : le temps de résidence moyen t_R du ligand sur la protéine réceptrice, exprimé en secondes.

$$t_R = \frac{1}{k_{off}} \quad [9]$$

Ce temps de résidence reflète le temps pendant lequel une molécule est active, car selon le principe d'Ehrlich : « une molécule n'est active que le temps durant lequel celle-ci est liée ».⁷

Si l'on considère qu'une enzyme catalyse une réaction de manière efficace, la vitesse de cette réaction est alors uniquement limitée par la vitesse de diffusion des substrats. Ces réactions contrôlées par la diffusion ont une constante de vitesse maximum k_{on} de l'ordre de $10^9 M^{-1}.s^{-1}$ selon Alberty, Hammes & Eigen (1958) limitées par la rencontre d'un substrat avec le récepteur.⁸ Cependant, Chou et coll. ont proposé un second modèle (1974) prenant en compte la capacité d'une enzyme à orienter à sa surface les substrats vers le site récepteur. Par ce modèle, les k_{on} peuvent atteindre une valeur maximale de l'ordre de $10^{10} M^{-1}.s^{-1}$.⁹

L'affinité et les constantes de vitesses dépendent tous deux du profil énergétique du phénomène d'interaction (Figure 2). A la différence de ΔG_b° , qui est calculée entre les états libres et liés et qui ne dépend que de ces deux états, les constantes de vitesses k_{on} et k_{off} sont fonction des énergies d'activation ΔG_{on}^\ddagger et ΔG_{off}^\ddagger calculées par la différence des énergies des états libres et liés avec l'état de transition (ET). Sur la Figure 2 on peut voir que si l'on déstabilise l'ET, seuls k_{on} et k_{off} seront diminués, sans altérer l'affinité (ligne rouge pleine) tandis que si l'on stabilise l'état lié, l'affinité sera

augmentée et k_{off} diminué tandis que k_{on} ne sera pas affecté (ligne rouge pointillée). A l'inverse si on altère l'énergie de l'état libre, k_{on} et K_D seront affectés mais pas k_{off} .

En pratique, il se trouve que pouvoir altérer deux de ces paramètres sans affecter le troisième est difficile. Cela montre qu'une optimisation basée uniquement sur la valeur du K_D peut masquer les variations des termes cinétiques, d'où l'importance dans notre étude de pouvoir vérifier, non pas uniquement la valeur de l'affinité des ligands pour la protéine d'intérêt, mais aussi leur capacité d'inhibition.⁵

I. 1. 4 La théorie de l'état de transition

Le concept d'état de transition est bien connu dans la description des réactions chimiques générales. On postule l'existence d'un état défini comme ayant une énergie maximale (Figure 2), un quasi-équilibre chimique entre les espèces réactives et le produit, sous forme d'un complexe activé dit de transition.



Lors d'une réaction chimique, les espèces réactives vont voir s'accroître leur énergie au fur et à mesure de l'avancement dans le chemin réactionnel jusqu'à atteindre ce maximum énergétique de transition. Il permet la formation du complexe activé appelé *état de transition*. C'est par la formation de ce complexe qu'est ensuite rendue possible la transformation en produit. La théorie de l'état de transition utilise l'équation d'Eyring afin de prédire la vitesse des réactions à partir des principes élémentaires de la thermodynamique.¹⁰ Cette théorie postule qu'à une température donnée, la vitesse de réaction ne dépend que de la concentration du complexe activé, en équilibre avec les réactifs.

Cet état de transition n'existe pas uniquement dans le cas d'interactions mettant en jeu des réactions chimiques. Dans le cas d'interactions PL simples il existe aussi des barrières de transitions énergétiques (transconformations, désolvatation, etc). La désolvatation par exemple, qui est un équilibre d'interaction simple sans réaction chimique, va posséder son état de transition (état libre non solvato), menant à l'état non solvato du complexe lié.

L'équilibre **10**, permet de définir la concentration en complexe activé à partir des concentrations en réactifs à l'aide de la constante d'équilibre :

$$[PL]^\ddagger = K^\ddagger [P][L] \quad [11]$$

La vitesse de réaction v , dépendante de la concentration en réactif peut alors être écrite sous forme de dérivée, telle que:

$$v = - \frac{d [P]}{dt} = k^\ddagger [PL]^\ddagger \quad [12]$$

Avec $[PL]^\ddagger$ la concentration en complexe activé, et k^\ddagger la constante de vitesse du passage à l'état de transition que l'on peut aussi développer à l'aide des termes de mécanique quantique :

$$k^\ddagger = \frac{k_B T}{h} \quad [13]$$

Avec k_B la constante de Boltzmann, T la température en Kelvin et h la constante de Planck.

En accord avec les lois de la thermodynamique classique, la constante d'équilibre peut aussi être reliée à la variation d'énergie pour la formation de l'état de transition à partir des réactifs, soit :

$$\Delta G^\ddagger = - RT \ln K^\ddagger \quad [14]$$

Soit, par passage à l'exponentielle :

$$K^\ddagger = e^{\frac{-\Delta G^\ddagger}{RT}} \quad [15]$$

Utilisant les équations **11** et **14**, la concentration de l'état activé pour l'interaction devient:

$$[PL]^\ddagger = e^{\frac{-\Delta G^\ddagger}{RT}} [P][L] \quad [16]$$

Ce qui permet d'exprimer la vitesse de réaction à l'aide des équations **12** et **13**, comme:

$$v = \frac{k_B T}{h} e^{\frac{-\Delta G^\ddagger}{RT}} [P][L] \quad [17]$$

Enfin, en reprenant l'expression de la vitesse donnée par la loi d'action de masse (équation **8**), on obtient l'expression de la valeur de la constante de vitesse k selon Eyring:

$$k = \frac{k_B T}{h} e^{\frac{-\Delta G^\ddagger}{RT}} \quad [18]$$

Cette équation démontre aussi que la valeur de l'énergie ΔG^\ddagger influe sur la vitesse de la réaction. Ainsi plus la valeur de ΔG^\ddagger est grande, et plus la réaction sera lente car l'énergie à apporter au système est plus importante pour induire la formation des produits. Il en résulte que si l'on veut atteindre l'équilibre chimique plus rapidement, il faut stabiliser cet état de transition en diminuant son énergie afin de favoriser le passage de la barrière énergétique menant à la formation du complexe PL.

I. 1. 5 Les différentes interactions mises en jeu

Les interactions mises en jeu dans le cas de l'interaction réversible entre une protéine et un ligand sont des interactions dites « faibles » et non covalentes comme les interactions de van der Waals (vdW), les interactions de charges coulombiennes ou les liaisons hydrogènes par exemple. Dans le cas de protéines et d'autres macromolécules biologiques ou assemblages supramoléculaires (acides nucléiques, membranes, etc), ces interactions se font par complémentarité, de nature chimique et structurale, au niveau d'un site particulier appelé le « site récepteur ».

- Complémentarité structurale

La complémentarité structurale entre une protéine et un ligand (ou un substrat pour une enzyme) a été définie par plusieurs modèles au cours du temps.^{11;12;13}

Le modèle de base appelé « le modèle clé-serrure », fut décrit par E. Fisher en 1894.¹¹ Ce modèle de reconnaissance est basé sur l'hypothèse d'une complémentarité de forme pré-existante entre le substrat et la cavité du site récepteur de la protéine. Il n'implique donc aucune modification de structure du récepteur ou du ligand (Figure 3). Dans ce modèle le récepteur et le ligand sont considérés comme des structures rigides.

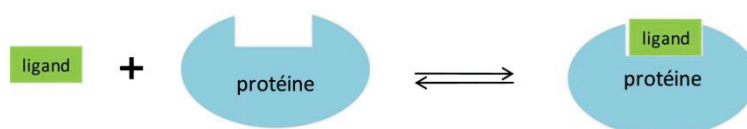


Figure 3. Schéma représentatif du modèle « clé-serrure ».

Modèle proposé par E. Fisher, basé sur la complémentarité enzyme-substrat dans lequel la protéine est assimilée à une structure fixe « serrure », et le ligand à une « clé » s'adaptant à cette serrure.

Suite à cette hypothèse, un second modèle a été amené par F. Karush en 1950¹⁴ puis D. Koshland en 1958.¹² Ce modèle est basé cette fois sur l'hypothèse que la structure de la protéine est modulable, et se déforme pour s'adapter à une petite molécule. Ce modèle de « l'ajustement induit » est un modèle dynamique, dans lequel une partie de l'énergie récupérée grâce aux interactions faibles formées entre les deux entités permet d'induire une déformation de la structure tridimensionnelle de la protéine, permettant la liaison du ligand dans le site récepteur (Figure 4). Cette déformation de la protéine peut conduire à une modification de son activité comme par exemple mener à une inhibition non compétitive, ou une régulation de l'allostérie.

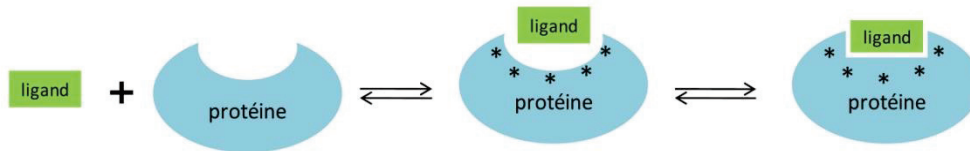


Figure 4. Schéma représentatif du modèle de « l'ajustement induit ».

Modèle dans lequel les interactions faibles (vdW, coulombiennes, liaisons H, électrostatiques) entre la protéine et le ligand (représentées par un *) induisent une déformation du site récepteur de la protéine, permettant alors la fixation du ligand.

Enfin, le dernier modèle est celui de la sélection conformationnelle, ou allostérie (Figure 5). Ce modèle postule l'existence d'un équilibre entre différentes conformations structurales de la protéine, avec la liaison préférentielle du ligand à l'une de ses formes. Ce modèle dynamique induit un déplacement de l'équilibre protéine-ligand vers la forme active de la protéine, et contrairement au modèle de l'ajustement induit, c'est l'interaction entre la protéine et le ligand qui va guider cet équilibre conformationnel vers la conformation active.¹⁵

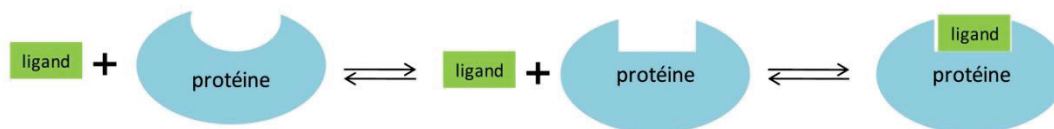


Figure 5. Schéma représentatif du modèle de la sélection conformationnelle.

Modèle dans lequel la protéine possède plusieurs conformations en équilibre, dont seulement l'une d'elles est favorable à la liaison du ligand dans le site récepteur.

- Complémentarité chimique

Les interactions pouvant exister au cœur du site récepteur, entre la protéine et son ligand sont des interactions de faible énergie ($< 10 \text{ kcal.mol}^{-1}$), non covalentes. Ces interactions sont à l'origine de la flexibilité et de la spécificité existant dans la plupart des procédés biologiques. Parmi ces interactions, on peut distinguer les liaisons hydrogène, les interactions coulombiennes, et les interactions de vdW.^{16;17}

Une liaison hydrogène est formée par l'interaction d'un atome d'hydrogène attaché à un atome électronégatif, avec un autre atome électronégatif (O, N ou S). Les liaisons hydrogènes sont les interactions les plus fréquemment observées entre deux biomolécules, car les protéines contiennent de nombreux groupes oxygènes et azotes, permettant leur formation. Les liaisons hydrogènes ont un grand rôle dans la stabilité structurale des macromolécules, car les acides aminés peuvent former

des liaisons hydrogènes intramoléculaires entre le groupement NH d'un résidu et le groupement C=O d'un autre résidu, formant la structure secondaire.

Lorsqu'un atome chargé positivement est mis en contact avec un atome chargé négativement, il est possible d'avoir une interaction de Coulomb (ou ionique). Cette interaction ionique est due à la force d'attraction électrostatique existante entre ces deux charges formant un dipôle électrique. Ces interactions coulombiennes coexistent avec les interactions apolaires et les liaisons hydrogènes et impliquent classiquement les acides aminés Arg, His, Lys (cations), et Asp ou Glu (anions).

D'autres acides aminés sont neutres et plutôt hydrophobes (Leu, Val, Ala, Ile, Phe, Trp, Tyr, Pro et Met) par la nature de leurs chaînes latérales. Deux groupements non polaires peuvent entrer en contact pour exclure le solvant (polaire). Il y a alors formation d'une interaction dite « hydrophobe », qui est en fait une combinaison d'interactions de vdW et d'un phénomène d'exclusion de solvant (H₂O). Dans ces interactions « hydrophobes » on retrouve aussi l'interaction π - π attractive pouvant être formée entre deux cycles aromatiques, appelé le « π -stacking », que l'on retrouve souvent dans la structure des protéines entre les acides aminés Phe, Tyr et Trp.^{2;17}

I. 2. Méthodes physico-chimiques pour l'étude des équilibres d'interactions

L'interaction d'un ligand avec une protéine va modifier la moyenne des environnements chimiques de chacun des partenaires. Lorsque le ligand est libre, cette moyenne est principalement due aux interactions fluctuantes avec le solvant, et éventuellement les sels dissous présents. Si la concentration en ligand devient plus importante, alors des interactions entre ligands pourront aussi avoir lieu et devenir significatives (solution non-idéale). Lorsque le ligand interagit avec la protéine, on peut distinguer des interactions dites non spécifiques souvent à la surface de celle-ci, combinées à des interactions plus spécifiques au sein du site d'interaction. Dans ce site d'interaction, le temps de résidence sera d'autant plus long que l'affinité sera grande. Nous allons passer brièvement en revue les différentes méthodes principalement utilisées pour définir les interactions entre ligand et protéines avant d'aborder les principaux avantages de la RMN en solution.³

I. 2. 1 Principales méthodes d'étude de l'interaction protéine-ligand

De nombreuses méthodes permettent la caractérisation de l'interaction protéine-ligand d'un point de vue structural (principalement la cristallographie aux rayons X ou la RMN), ou d'un point de vue cinétique et thermodynamique avec la résonance plasmonique de surface (SPR),¹⁸ les méthodes calorimétriques (ITC ou calorimétrie différentielle),^{19; 20; 21} les méthodes séparatives (dialyse,

ultrafiltration, ultracentrifugation, chromatographie, électrophorèse),^{22; 23} le dichroïsme circulaire^{24;25} ou bien encore la spectrométrie de masse (ionisation nano-electrospray). L'ensemble de ces techniques possède des caractéristiques qui peuvent être parfois complémentaires dans l'étude de l'affinité.

La cristallographie aux rayons X est par exemple une méthode de choix quand il s'agit d'obtenir des informations structurales en général. Cependant elle requiert l'obtention de mono-cristaux de bonne qualité et des complexes PL définis.²⁶ Les cristaux de complexes peuvent être obtenus par trempage des cristaux de la protéine dans des solutions contenant le ligand. Celui-ci diffuse au travers du réseau cristallin via les canaux de solvant. Si la méthode par trempage n'est pas efficace, alors il est possible de réaliser des expériences de co-cristallisation d'une solution de protéine et de ligand. L'analyse des cartes de densité électronique permet ensuite l'identification des conformations liées du ligand, et la résolution de la structure complète du complexe PL.²⁷ La cristallographie est une méthode qui demande une concentration importante en protéine, et qui ne renseigne que qualitativement sur la dynamique des molécules au sein du réseau cristallin. L'ordre du cristal organise les molécules en son sein, et aucune information quant à l'affinité réelle ne peut être obtenue.

La SPR et l'ITC sont deux techniques qui ont été largement utilisées dans la littérature récente pour la détermination de l'affinité PL via la détermination des constantes thermodynamiques et cinétiques.

- Résonance plasmonique de surface (SPR)

La technique SPR permet d'observer la cinétique d'association (k_{on})/dissociation (k_{off}). Cette technique est applicable aux couples protéine/ADN, protéine/protéine mais aussi protéine/ligand. Elle donne ainsi accès à la constante d'équilibre d'affinité du complexe (relation 6).

La SPR est basée sur la détection optique des variations de masse sur une surface d'or placée sur une puce de détection en verre. La protéine doit être immobilisée par un greffage chimique sur cette surface par une couche de Dextran (polymère ramifié de glucose), et le ligand est injecté par un système micro fluidique dans un flux continu à la surface du capteur (Figure 6). Le système de détection consiste à mesurer la variation de l'indice de réfraction au voisinage de l'interface, lorsque le ligand est en liaison avec son récepteur.²⁸ La SPR a un inconvénient majeur : l'immobilisation de la protéine d'intérêt pouvant mener à des artefacts.

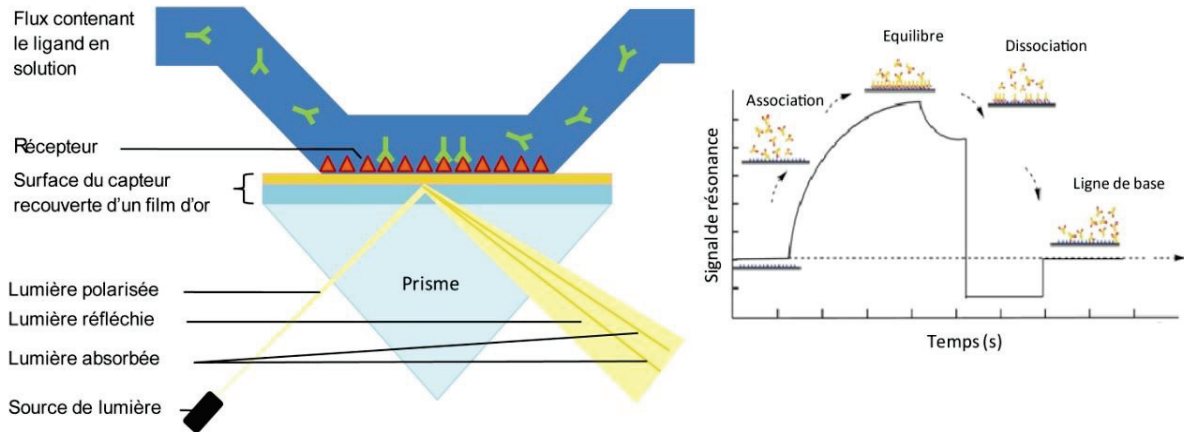


Figure 6. Schéma du principe de détection par la technique SPR.

La méthode SPR, où le récepteur correspond à la protéine greffée sur une matrice (dextran). Détection de la variation de l'indice de réfraction entre la phase contenant la protéine et celle contenant la molécule ligand. La variation de l'angle de résonance est suivie en temps réel et peut être tracé comme signal de résonance en fonction du temps.

- Titration calorimétrique isotherme (ITC)

L'ITC repose sur la mesure directe de la chaleur libérée ou absorbée durant l'interaction (Figure 7) et permet la mesure des grandeurs thermodynamiques d'enthalpie ΔH , de l'énergie libre ΔG (par détermination du K_D), et donc de l'entropie ΔS .

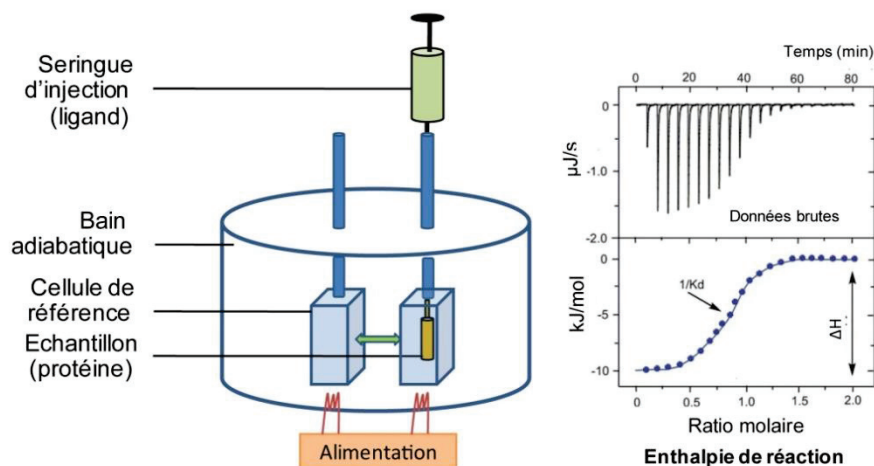


Figure 7. Schéma du principe de détection par la technique ITC.

La méthode de l'ITC est basée sur la mesure de la chaleur libérée lors de la fixation d'un ligand sur la protéine cible. Lors d'une titration le thermogramme obtenu représente la chaleur libérée après chaque injection, en fonction du temps. Il est ensuite possible de représenter la chaleur libérée après chaque ajout, en fonction du ratio molaire $[L]/[P]$ pour obtenir la courbe expérimentale (points bleus) à laquelle on fait correspondre la courbe théorique du modèle considéré (selon le nombre de sites de liaison) pour déterminer K_D .

De plus, elle ne nécessite aucune modification des partenaires impliqués dans la liaison et permet la mesure de leur affinité dans leur état natif.¹⁹ C'est à ce jour la seule méthode permettant la mesure de la variation d'enthalpie ΔH . Cependant, c'est une technique macroscopique globale qui n'apporte aucune information structurale sur l'interaction.^{19;21}

Nous allons voir que la RMN, en permettant une analyse à la fois structurale et thermodynamique, est une méthode très riche pour l'étude d'équilibres d'interaction PL.

I. 2. 2 Utilisation de la résonance magnétique nucléaire (RMN) en solution

La RMN permet de détecter des interactions protéine-ligand pour un large domaine d'affinité et de taille de protéines et rend aussi possible la détermination qualitative et quantitative de l'affinité PL avec des informations à la fois structurales et dynamiques. En effet les modifications subies par chacun des deux partenaires de l'interaction auront un effet direct sur les observables RMN. L'étude de ces différents observables, rendue possible grâce à une palette très large d'expériences, va permettre d'apporter des informations diverses. Le déplacement chimique isotrope δ , par exemple, va donner des informations sur l'interaction PL. Il en est de même des expériences de NOE transférés et de STD qui sont basées sur un transfert d'aimantation intra ou inter moléculaire entre la protéine et le ligand, ou encore les expériences de WaterLogsy basées sur la différence de relaxation entre les molécules d'eau libres et liées en solution.²⁸ Les expériences basées sur la différence des vitesses de relaxations transversales (R_2) et longitudinales (R_1) ainsi que les expériences NOESY et l'étude des couplages dipolaires résiduels (RDC) seront, elles, utilisées pour observer des modifications structurales et dynamiques au niveau du récepteur. De par sa résolution, la RMN est considérée comme la méthode la plus polyvalente pour l'étude de l'affinité.^{2;29;30}

- Le déplacement chimique en RMN

La spectroscopie RMN est une technique qui permet l'observation de signaux de radiofréquence provenant de noyaux atomiques de spin non nul. Lorsqu'ils sont placés dans un champ magnétique uniforme, la radiofréquence à laquelle sont observés les signaux, est définie comme la fréquence de Larmor ν_0 :

$$\nu_0 = \frac{\gamma B_0}{2\pi} \quad [19]$$

Où B_0 représente la valeur du champ magnétique de l'appareil exprimé en Tesla, et γ le rapport gyromagnétique du noyau observé exprimé en $\text{rad}\cdot\text{s}^{-1}\cdot\text{T}^{-1}$. Cependant, cette fréquence de résonance est modulée par l'environnement chimique. La différence de fréquence induite par ces environnements, est appelé le déplacement chimique δ , qui est de l'ordre de la partie par million

(ppm) de ν_0 . En solution où les molécules diffusent librement, le déplacement chimique représente une valeur moyenne de toutes les orientations possibles en solution. Cet observable RMN intégrant l'ensemble des orientations possibles peut être plus particulièrement mis en parallèle des simulations de dynamique moléculaire qui seront déployées dans la suite de notre étude.

En fonction de la cinétique de l'échange, les spectres seront différemment affectés. Dans le cas d'un échange rapide, seul un δ moyen sera observé sous forme d'une raie fine, tandis que dans le cas d'un échange lent entre deux formes libres et liées, deux δ seront observés : un δ pour la forme libre et un δ pour la forme liée (Figure 8). Enfin, si l'échange se fait à vitesse intermédiaire les raies observées seront élargies sous l'effet de l'interaction, et avec l'augmentation de la vitesse on atteindra un régime de coalescence où l'on ne distinguera qu'une seule raie très large.

Nous avons dans ce projet essentiellement utilisé l'observable déplacement chimique, ainsi que le transfert de saturation.

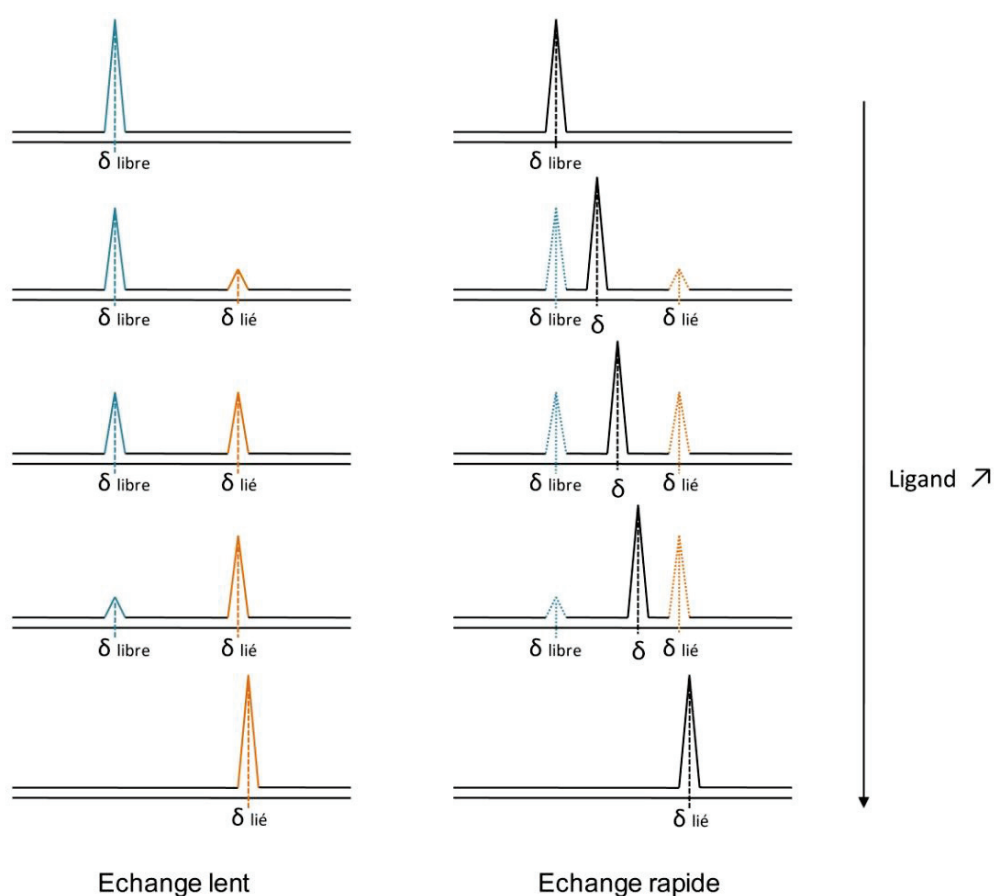


Figure 8. Influence de la vitesse de l'échange sur les signaux RMN.

Le déplacement chimique étant sensible à l'environnement chimique de l'atome considéré, il y aura modification de sa valeur lors de l'interaction avec un ligand. On a donc δ libre (bleu) différent de δ lié (orange). Selon la cinétique de l'échange, le spectre RMN va être modifié. Dans le cas d'un échange lent, deux δ seront observés : un δ pour la forme libre et un δ pour la forme liée. Dans le cas d'un équilibre rapide, seul le δ moyen (noir) sera observé.

I. 2. 3 Observation des signaux RMN provenant de la protéine : l'expérience HSQC

L'une des méthodes les plus utilisées pour l'étude de l'affinité protéine-ligand en RMN est l'observation des signaux de la protéine réceptrice par l'expérience 2D HSQC (*Heteronuclear Single Quantum Coherence*).³¹ Dans cette expérience les δ ^1H sont corrélés par les couplages spin-spin $^1J_{\text{HX}}$ avec δ ^{15}N ou δ ^{13}C .

Au travers des δ ^1H et ^{15}N on peut aisément détecter des variations de déplacement chimique (CSP) induites par l'interaction avec la molécule ligand en fonction de sa concentration par rapport à une concentration en protéine fixe. L'HSQC implique un transfert de cohérence du proton vers l'hétéroatome considéré, via une séquence d'impulsion INEPT (*Insensitive Nuclei Enhanced by Polarization Transfer*).³² Très sensible et rapide à acquérir, elle permet d'obtenir une signature caractéristique de la protéine dans des conditions expérimentales données (Figure 9). Cette expérience va non seulement rendre possible l'identification des molécules qui pourraient se lier à la protéine cible (présence de signaux perturbés par la liaison), mais aussi de localiser un supposé site d'interaction en identifiant les pics de corrélation des acides aminés perturbés par la présence du ligand sur le spectre 2D, grâce à l'attribution préalable des résonances (^1H et ^{15}N dans notre cas) de la protéine. Cependant, des modifications distantes du site de liaison sont aussi possibles en cas de changements de conformation de la protéine (modèle de l'ajustement induit de Koshland). De plus, en titrant la protéine par des ajouts croissants de ligand il sera possible de déterminer des valeurs de K_D apparents (voir partie IV.2 du chapitre II).

Pour ce type d'expérience, il existe néanmoins certaines contraintes et limitations. Tout d'abord, l'abondance naturelle de l'azote ^{15}N est telle (0.4 %) qu'il est nécessaire de marquer isotopiquement la protéine d'intérêt afin de pouvoir l'étudier. En effet l'azote ^{14}N a un spin entier et un moment quadripolaire ce qui rend la détection en général impossible, on cherche donc à observer le noyau ^{15}N qui possède un spin $\frac{1}{2}$. De plus, il est difficile de pouvoir analyser des protéines de haut poids moléculaires (MW), on limitera donc son utilisation à des systèmes de MW inférieur à 50 kDa, et le temps d'expérience peut atteindre plusieurs heures selon la concentration de la protéine disponible.

I. 2. 4 Observation des signaux RMN provenant du ligand : l'expérience STD

La détection de l'affinité en RMN peut aussi être réalisée indirectement par l'étude des spectres de ligands libres. Cette détection indirecte est intéressante pour détecter l'affinité en présence d'édifices macromoléculaires dont les caractéristiques de relaxation rendent inefficaces les expériences HSQC ou TROSY.

Le principe général de l'expérience STD est l'irradiation sélective d'un signal spécifique de la protéine dont la saturation sera transférée à l'ensemble de la protéine puis au ligand (Figure 10). Elle ne nécessite qu'une faible concentration en protéine (entre 1 μM et 20 μM selon la taille), et un excès de ligand (rapport de concentration protéine:ligand entre 1:100 et 1:1000).³³ Le STD peut aussi être utilisé sur un ligand particulier dont on fait varier progressivement la concentration. A chaque valeur de concentration en ligand, pour une concentration en protéine constante, on effectue un spectre de référence 1D ^1H et un spectre de différence STD. En mesurant ensuite l'intégrale des signaux du ligand sur chacun des spectres, on peut calculer un facteur d'amplification STD (f_{STD}) :

$$f_{STD} = \frac{I_{STD}}{I_{1D}} \cdot \frac{[L]}{[P]} \quad [20]$$

Avec I_{STD} et I_0 l'aire sous les pics sur les spectres STD et de référence respectivement.

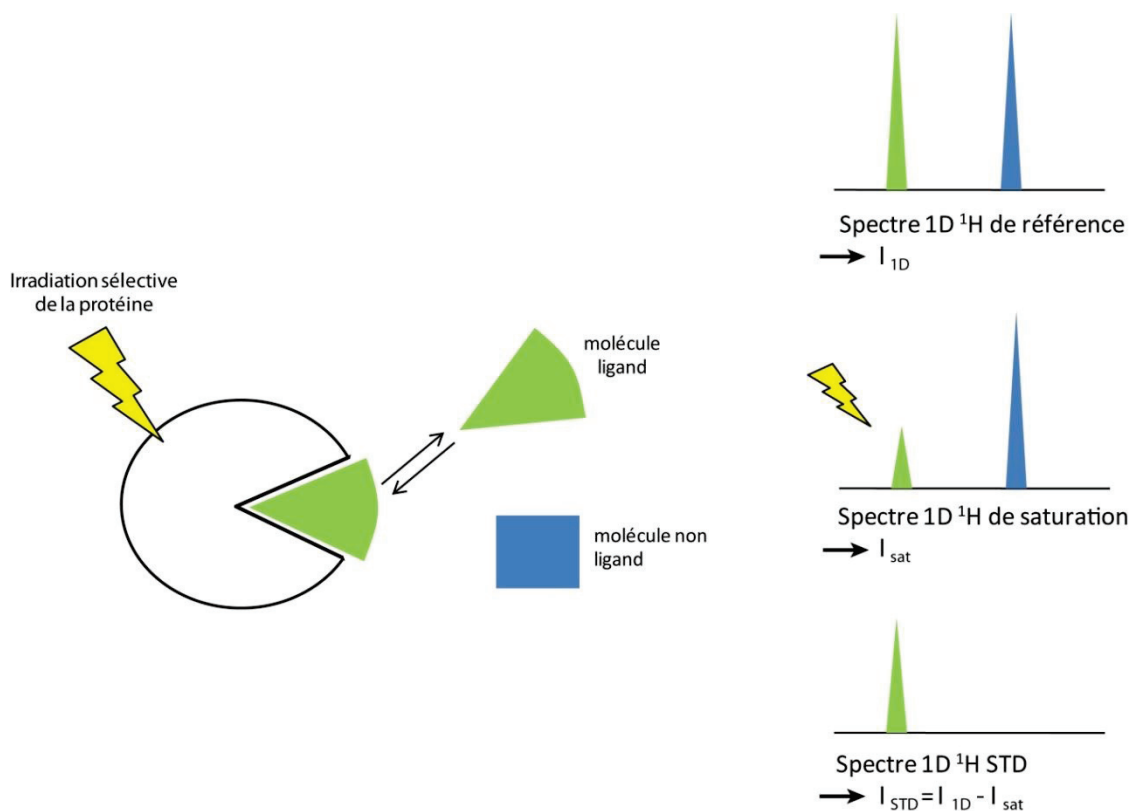


Figure 10. Principe de l'expérience RMN 1D ^1H de transfert de saturation (STD).

Visualisation des signaux RMN d'une molécule ligand (vert) et d'une molécule non ligand (bleu) pour les différents spectres acquis : le spectre 1D de référence, le spectre acquis après irradiation sélective d'un des signaux de la protéine, et le spectre obtenu par différence. Cartographie d'épitopes : Visualisation de la différence d'intensité entre deux signaux du ligand sur le spectre de saturation, permettant de définir une conformation prédominante dans laquelle le ligand est lié au récepteur. Le proton le plus impliqué dans l'interaction verra son signal plus fortement diminué par l'irradiation sélective de la protéine.

Lorsque l'on représente f_{STD} en fonction de la concentration en ligand, on peut ajuster non linéairement des données expérimentales avec une hyperbole. Cet ajustement permet d'obtenir une estimation du K_D apparent du complexe, mais aussi l'implication des différents protons du ligand dans l'interaction PL. Ainsi, le proton ayant le plus fort facteur d'amplification STD sera considéré comme le proton le plus en contact avec la protéine. Cette analyse qu'on appelle la « cartographie d'épitopes » (ou *epitope mapping*) permet de supposer des conformations liées préférentielles.³⁴

I. 3. Modélisation utilisant la dynamique moléculaire

En regard de la RMN en particulier, il existe aussi des moyens théoriques pour la modélisation des équilibres d'interactions, la modélisation pouvant devenir dans certains cas prédictive et quantitative. L'évolution de la puissance de calcul des ordinateurs au cours des dernières décennies a permis l'émergence de nouveaux domaines de modélisation, en étendant en particulier les échelles de temps de la simulation de phénomènes dynamiques tels que la dynamique moléculaire.

I. 3. 1 Les bases de la modélisation moléculaire

La modélisation moléculaire permet, grâce à des méthodes numériques, de mieux comprendre les interactions entre biomolécules en simulant leur comportement à différentes échelles de temps. Il serait ainsi ultimement possible de relier l'observation expérimentale macroscopique à son aspect microscopique.³⁵

- La mécanique quantique

Afin de pouvoir relier les propriétés d'un système microscopique à celles de la phase macroscopique, il est avant tout nécessaire de restreindre le nombre de particules simulées, car en toute rigueur l'étude d'un système complexe moléculaire nécessiterait la résolution de l'équation de Schrödinger :

$$H_0 \cdot \psi_0 = E_0 \cdot \psi_0 \quad [21]$$

Avec H_0 l'opérateur hamiltonien du système, E_0 son énergie, et $\psi_0 (r_1, \dots, r_N)$ une fonction d'onde décrivant l'état du système, avec N le nombre de particules.³⁶

Or, dans le cas général d'un ensemble polyatomique et électronique, cette équation n'admet aucune solution analytique exacte. C'est pour cette raison que des approximations ont été faites, afin d'obtenir des énergies E et fonctions d'onde ψ qui s'approchent le plus possible de E_0 et ψ_0 . La plus importante d'entre elles a été définie par Born et Oppenheimer en 1927.³⁷ Elle est fondée sur l'hypothèse que l'on peut découpler le mouvement des électrons et des noyaux, qui ont une masse plus importante. Le mouvement des noyaux est ainsi négligé, ce qui revient à considérer que les électrons se déplacent dans un champ de noyaux fixes dans l'espace. Ainsi, l'énergie du système moléculaire considéré ne dépend plus que de la position des noyaux, pour lesquels il est possible d'utiliser une description newtonienne du mouvement. La restriction du nombre de particules simulées est effectuée en ayant recours aux conditions de limites périodiques, qui consistent à répliquer un ensemble fini de particules dans une boîte (cellule de simulation), selon les 3 directions de l'espace (Figure 11).³⁸

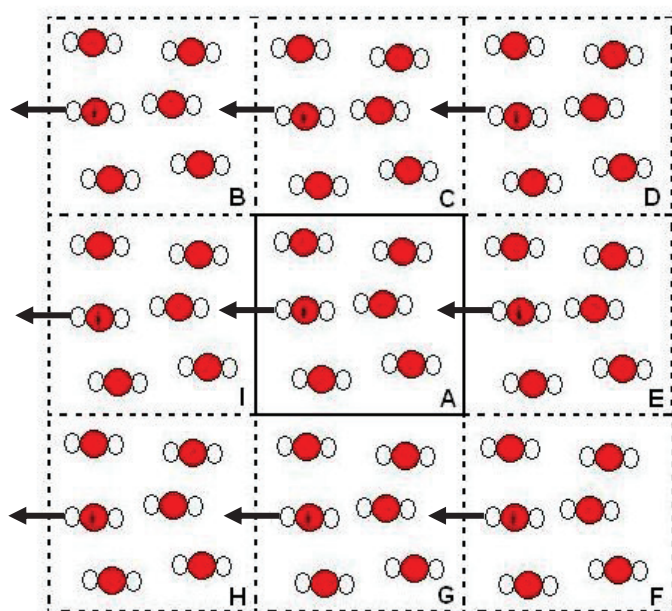


Figure 11. Vue bidimensionnelle d'une cellule de simulation répliquée dans les trois dimensions de l'espace. Avec les conditions aux limites périodiques, lorsque la particule i quitte la boîte centrale A, ses images situées dans les cellules voisines vont se déplacer de manière similaire.

Le système généré possède alors un caractère pseudo-infini, ce qui induit l'utilisation d'autres approximations. En effet chaque particule interagit alors avec les autres particules de la boîte, mais aussi avec ses images virtuelles.³⁸ L'emploi de méthodes de type somme de réseau, telle que la sommation d'Ewald, permet de calculer le potentiel électrostatique total en répartissant le calcul entre un espace réel et un espace réciproque (par transformée de Fourier). Dans notre cas, la méthode utilisée est une version simplifiée de la sommation d'Ewald, appelée « *Particle Mesh Ewald* » (PME).³⁹

- Les champs de forces de la mécanique moléculaire classique

Toutes les tâches intervenant en modélisation moléculaire nécessitent le choix et la définition des champs de forces à utiliser.

Un champ de forces est la somme des termes caractérisant les interactions entre les différents atomes d'une molécule. Il régit donc l'énergie globale d'un système moléculaire. Celui-ci est défini par un ensemble de fonctions empiriques et paramétriques décrivant l'énergie potentielle du système de particules considéré.³⁶ Le champ de force comprend les termes liés à deux types essentiels d'interactions : les interactions covalentes dépendantes de la géométrie de la molécule, et les interactions non-covalentes qui ne concernent que les atomes non liés entre eux.³⁶

Les termes représentant les interactions covalentes sont principalement divisés en 2 catégories : les déformations de longueur de liaison, les distorsions des angles de valence, et les termes diédraux (torsions). Dans le terme exprimant les interactions non-liantes, on trouve la contribution coulombienne représentant les interactions électrostatiques, et les termes de van der Waals (Figure 12). Les champs de forces classiquement utilisés dans le cas de systèmes biologiques sont AMBER (*Assisted Model Building with Energy Refinement*), CHARMM (*Chemistry at HARvard Macromolecular Mechanics*), ou GROMOS (*GRONingen MOlecular Simulation Package*). Ils sont développés de façon communautaire depuis plusieurs dizaines d'années.^{40; 41; 42}

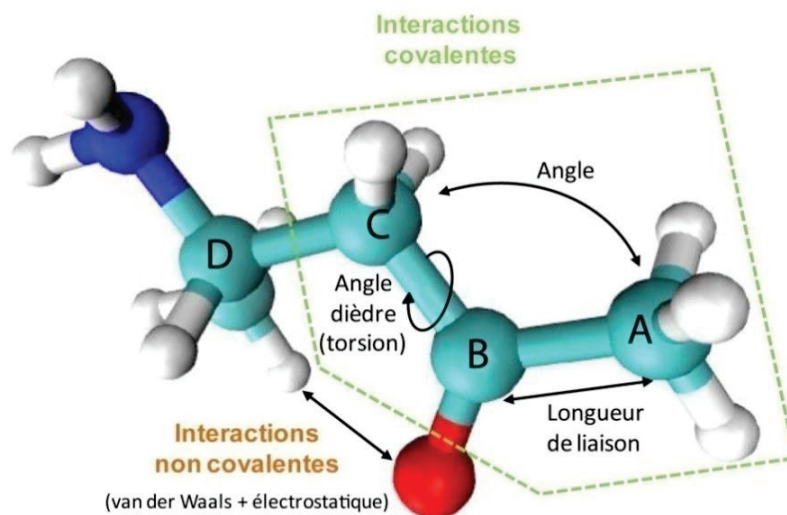


Figure 12. Visualisation des différentes composantes des termes énergétiques prises en compte dans le potentiel d'une molécule en mécanique moléculaire pour la définition des champs de force .

Encadré vert : interactions covalentes comportant les longueurs de liaison, les valeurs d'angles et de torsions. En orange, les interactions non liantes comprenant les interactions de faible énergie de type van der Waals et interactions coulombiennes pouvant exister entre une protéine et un ligand.

I. 3. 2 La dynamique moléculaire

Les protéines sont des entités qui possèdent une dynamique intrinsèque de par la flexibilité des acides aminés qui les composent. Leur dynamique concerne aussi leur diffusion au sein de leur milieu. De même, les ligands possèdent eux aussi leur propre dynamique, ainsi que le solvant.⁴³ La dynamique moléculaire (MD) permet de simuler le comportement d'un système moléculaire au cours du temps, par intégration des équations du mouvement de Newton, grâce à l'approximation de Born-Oppenheimer (voir I.3.1).

- Les équations du mouvement

La dynamique moléculaire consiste donc à résoudre les équations du mouvement définies par Newton :

$$m_i \cdot a_i(t) = F_i(t) \quad [22]$$

Où F_i est la force exercée sur la particule i , m_i sa masse, et $a_i = \frac{dv_i}{dt}$ son accélération.

La force sur un atome est alors obtenue à partir des dérivées de l'énergie potentielle V par rapport aux coordonnées de l'atome r_i :

$$m_i \frac{\partial^2 r_i(t)}{\partial t^2} = - \frac{\partial V}{\partial r_i} \quad [23]$$

Les équations différentielles **23** du mouvement sont ensuite résolues numériquement par des méthodes itératives. A partir des coordonnées et des vitesses initiales d'un atome au temps t , les positions et vitesses de cet atome sont calculées pour le temps $t + \Delta t$. Le pas Δt de l'intégration définit l'intervalle de temps durant lequel les forces exercées sur les atomes restent constantes. C'est pour cette raison que plus un système est grand, plus l'évaluation de l'énergie de l'ensemble des atomes est coûteuse en temps de calcul et en mémoire. Ce pas d'intégration est à définir en fonction du système et donc des types de mouvements que l'on cherche à modéliser. Pour des molécules rigides, un pas assez grand de 5 fs sera approprié, mais ne le sera pas pour une molécule plus flexible dont les mouvements seront plus importants. Dans ce cas il faut diminuer le pas d'intégration, typiquement de l'ordre de 1 fs.

- Thermostat et barostat

En dynamique moléculaire classique, le nombre de particules N , le volume V et l'énergie totale du système E sont conservés ce qui permet de définir en mécanique statistique un ensemble micro-canonique. Cependant, les approximations faites peuvent mener à des instabilités et donc à

des variations d'énergie ou de température.⁴⁴ Afin de modéliser l'environnement du système de manière efficace, il est important de pouvoir réaliser des simulations de dynamique moléculaire avec contrôle de la température (thermostat) et/ou de la pression (barostat) comme grandeurs indépendantes. Cela permet d'effectuer des simulations dans des ensembles thermodynamiques spécifiques différents : micro-canonique (NVE), canoniques (NVT) ou isotherme-isobare (NPT) par exemple, dans lesquels l'ensemble des atomes possèdent le même état macroscopique dépendant des paramètres thermodynamiques classiques. Dans ces ensembles, le nombre de particules est toujours gardé constant.³⁸

La température étant liée à l'énergie cinétique du système, il est possible de la contrôler en ajustant les vitesses des atomes du système soit pas à pas (thermostat d'Andersen), soit avec un algorithme qui régule l'énergie cinétique des molécules en fonction de la différence de température (thermostat de Berendsen). Une dernière technique est d'introduire un degré de liberté supplémentaire directement dans les équations du mouvement (thermostat de Langevin ou de Nosé-Hoover).

La pression du système quant à elle, est liée à son volume. Il est donc possible de contrôler la pression en faisant varier le volume de la boîte de simulation, soit en ajustant la position des particules les unes par rapport aux autres, soit en utilisant un algorithme qui relie les fluctuations de pression au volume et maintient la pression constante par ajustement des coordonnées atomiques de la boîte de simulation (barostat de Berendsen). Enfin, de la même manière que dans le cas de la régulation de température, il est possible d'ajouter une variable externe caractérisant le volume de la boîte de simulation dans les équations du mouvement (barostat d'Andersen).

I. 3. 3 Prédiction de l'affinité protéine-ligand par MD

Il est possible, en utilisant la dynamique moléculaire, d'avoir une estimation qualitative et théorique de l'énergie libre d'interaction. Or il est nécessaire de pouvoir connaître les modifications d'énergie induites par un phénomène biologique pour le comprendre de manière complète. La simulation de ces changements énergétiques donne accès aux affinités simulées de complexes PL, ce qui permet de modéliser et de visualiser de manière microscopique et statistique les mécanismes d'interaction ayant lieu entre les deux partenaires biologiques au niveau de la molécule.^{45;46} De plus, cela pourrait avoir un impact sur la manière de concevoir des nouvelles molécules inhibitrices, et apporter de nouveaux outils théoriques pour la conception de médicaments dans le domaine médical. C'est donc un défi auquel de nombreux chimistes théoriciens se sont confrontés.^{47;48}

L'énergie libre peut être exprimée de manière absolue, ou de manière relative par rapport à un état de référence. Comme décrit par l'équation 5, nous avons choisi d'exprimer cette énergie dans ses

conditions standards (par mole), de manière absolue : ΔG°_b . Plusieurs stratégies ont émergées pour le calcul des énergies libres de liaison que l'on peut séparer en deux catégories : les méthodes dites « *end-points* » et les méthodes « *free-energy pathway* ». ⁴⁹

- Méthode « *end-points* » : différence calculée entre 2 points

Pour cette méthode, un échantillonnage séparé des conformations libres et liées du complexe est réalisé. Les principales contributions physiques entrant dans le calcul de l'énergie libre sont déterminées pour chacune de ces conformations distinctes (l'énergie interne, l'entropie, l'énergie de solvatation). Puis, l'énergie libre de liaison est ensuite calculée par différence des composantes dans l'état libre et l'état lié du complexe. Cette manière de calculer l'énergie est utilisée dans de nombreuses méthodologies de MD telles que la méthode d'énergie d'interaction linéaire (LIE), l'approximation de réponse linéaire (LRA) ou la mécanique moléculaire d'aire de surface de Poisson-Boltzmann (MM-PBSA) par exemple. Mais elle ne permet pas de mettre en évidence les possibles états de transition par lesquels le système passe. ^{48;50;51}

- Méthode « *free-energy pathway* » : somme des contributions le long du chemin

Au contraire de la méthode précédente, cette méthode permet un échantillonnage des conformations existantes entre les formes libres et liées du complexe. Le ligand est graduellement séparé de la protéine, et l'énergie libre de liaison est obtenue par sommation de toutes les contributions énergétiques existantes entre l'état initial et l'état final. C'est l'approche de calcul utilisée dans la méthode d'intégration thermodynamique (TI) et la méthode de perturbation d'énergie libre (FEP pour *Free Energy Perturbation*). ^{45;52} Ce type de détermination estime de manière plus précise la valeur de ΔG°_b , car elle décompose le terme énergétique en plusieurs contributions (répulsion des noyaux, dispersion de van der Waals, interactions électrostatiques) mais elle demande une puissance de calcul plus importante. ⁴⁶ Nous verrons par la suite qu'avec les progrès de ces dernières années en matière de processeurs de calcul, ces techniques ont été de plus en plus utilisées dans un contexte de conception de molécules bioactives.

- Détermination des constantes cinétiques

Comme nous l'avons vu précédemment, la valeur de l'affinité est aussi reliée aux valeurs des constantes cinétiques d'association et de dissociation. Une troisième méthode est donc de déterminer par MD les valeurs de k_{on} et k_{off} , afin d'en déduire une affinité théorique.

Pour cela, il faut d'abord simuler l'étape de dissociation en partant du complexe lié et observer le ligand se dissocier afin de déterminer k_{off} . Il faut ensuite partir des espèces libres et observer l'étape

de liaison pour la détermination de k_{on} . Dans les deux cas un grand nombre de simulations sont nécessaires afin d'obtenir une valeur moyenne de la constante de vitesse d'association ou de dissociation, du fait du caractère aléatoire de la MD.^{46;53}

- Effet du solvant

Certains paramètres sont à adapter en fonction du système et ce quelque soit la méthode utilisée comme la contribution du solvant. Le rôle de l'eau est crucial dans de nombreux procédés biologiques, et l'estimation de l'énergie d'interaction dans notre cas va fortement dépendre de la bonne description des contributions du solvant.⁵⁴

Des approximations ont été faites pour considérer les molécules d'eau comme un milieu continu homogène du fait du faible temps nécessaire à ces molécules pour réagir à une perturbation et revenir à leur état d'équilibre. Ces méthodes dites en « solvant implicite » induisent un effet moyenné du solvant, pour des temps de calcul fortement diminués. Cependant, l'inconvénient majeur réside dans la diffusion lente des molécules d'eau présentes dans des cavités peu accessibles, l'effet des molécules d'eau de manière indépendante sur les liaisons hydrogènes formées, et les effets d'hydrophobicité en général.⁵¹

Les études en « solvant explicite » considèrent les molécules de solvant par leurs coordonnées réelles et certains degrés de liberté. On distingue ces modèles (type TIPXP) de par le nombre de sites X employés pour définir les atomes de solvant (Figure 13). Le modèle TIP3P par exemple, très couramment utilisé, considère une molécule d'eau comme ayant 3 sites d'interactions, correspondant aux 3 atomes de la molécule d'eau, à qui sont attribués chacun une charge ponctuelle.⁵⁵ Mais des modèles plus complexes à 5 ou 6 sites existent au prix d'un temps de calcul prolongé.⁵⁶

I. 3. 4 De la dynamique moléculaire à la funnel-métadynamique

Les développements récents de la MD avaient pour motivation d'obtenir des résultats plus précis par rapport à l'expérience. De même il a été recherché une optimisation du temps de simulation et un accès à des échelles de temps plus longues afin de pouvoir étudier des événements ayant lieu sur des échelles de temps plus grandes que la centaine de nanosecondes.

De cette volonté, et grâce à l'émergence de meilleures performances numériques (dont les processeurs graphiques GPU)(Figure 14) ou de processeurs dédiés (ANTON),⁵⁷ sont apparues de nouvelles méthodologies comme la dynamique moléculaire à haut débit (HTMD)^{58; 59} ou la dynamique moléculaire accélérée (aMD)⁶⁰ par exemple, mais aussi la métadynamique.^{61; 62} Le

concept de métadynamique est dérivé de la dynamique moléculaire classique, à la différence qu'un biais va être appliqué au système modélisé, sous forme de contraintes.

La métadynamique a été particulièrement étudiée dans l'équipe du Pr. Parrinello (ETH Zurich et Université de Lugano). Ce type de MD permet d'accéder à des échelles de temps plus importantes, et donc l'observation de phénomènes « rares » tels que la liaison d'un ligand à son récepteur ou la dissociation d'un ligand à haute affinité.⁶³

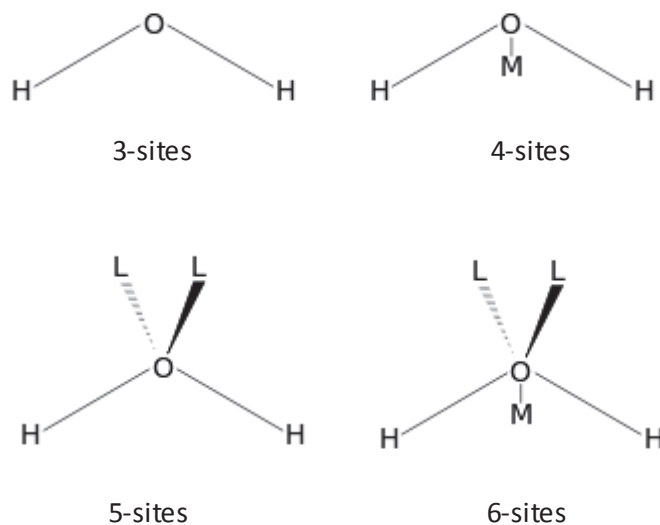


Figure 13. Représentation graphique des différents modèles de solvant explicite définis selon leur géométrie et le nombre de sites utilisés.

TIP3P : 3 sites d'interaction correspondant aux 3 atomes de la molécule d'eau ;

TIP4P : 4 sites, les 3 atomes de la molécule plus un pseudo-atome M (représentant la charge négative de l'oxygène) ;

TIP5P : 5 sites, les 3 atomes de la molécule plus 2 atomes fictifs L (doublets non-liants de l'oxygène) ;

TIP6P : modèle le plus complexe à 6 sites combinant 2 atomes fictifs et un pseudo-atome en plus des 3 atomes de la molécule d'eau.

- Le principe de la métadynamique

La métadynamique est une technique de dynamique moléculaire qui permet d'améliorer l'échantillonnage des simulations, et de pouvoir cartographier l'énergie du système en fonction de certains degrés de liberté du système qui sont appelés les *variables collectives* (CV). Ce sont ces variables collectives qui vont être biaisées afin d'accélérer l'échantillonnage.⁶¹ Le biais introduit est un potentiel externe dépendant du chemin déjà parcouru lors de la simulation, qui est fonction d'une ou plusieurs CV.

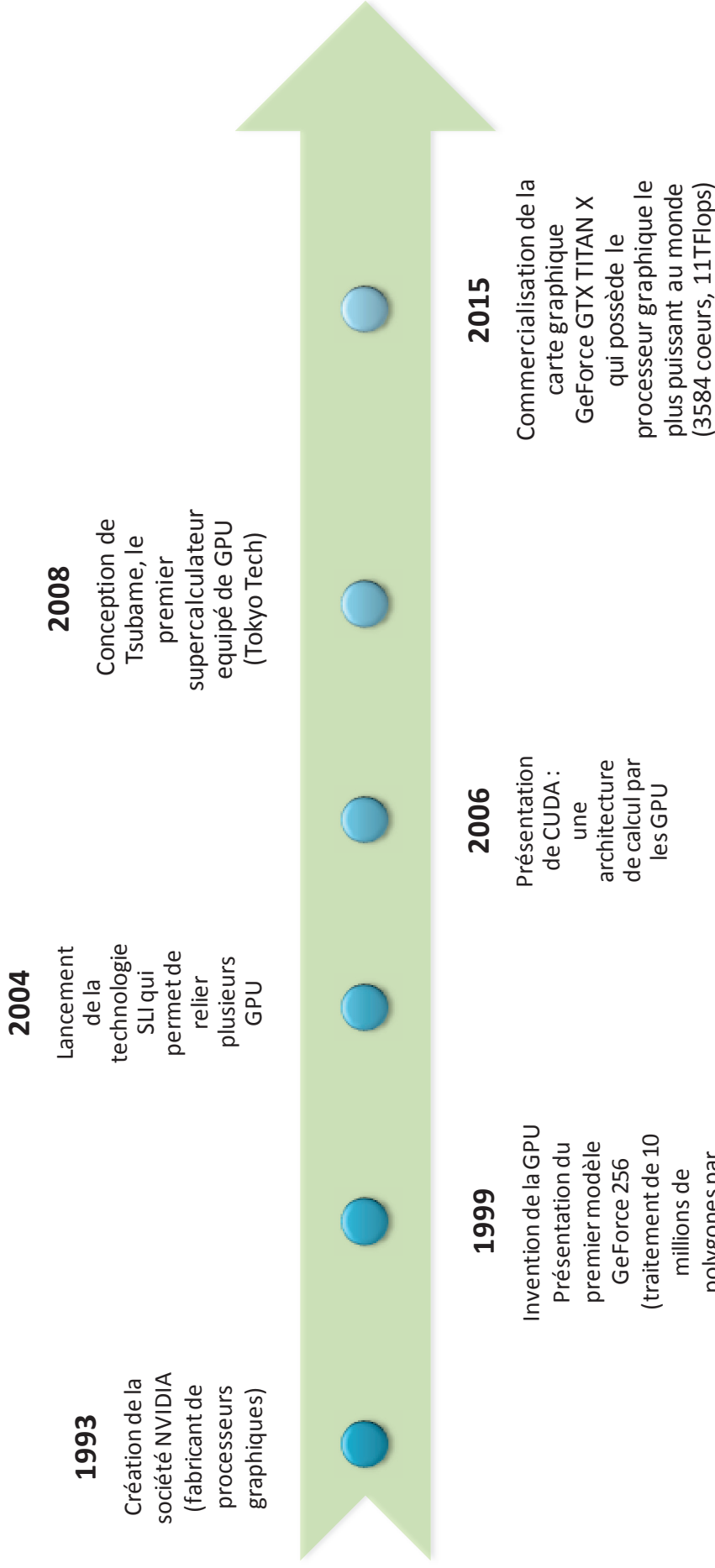


Figure 14. Illustration schématique des progrès des processeurs graphiques à travers l'essor de la société NVIDIA®.
SLI : Scalable Link Interface / GPU : Graphic Processor Unit / CUDA : Compute Unified Device Architecture

Ce potentiel est construit comme une somme de gaussiennes (appelé potentiel de Hills), et déposé le long de la trajectoire dans l'espace suivi par les CVs, afin de décourager le système de revisiter les conformations qui ont déjà été échantillonnées (Figure 15).⁶⁴

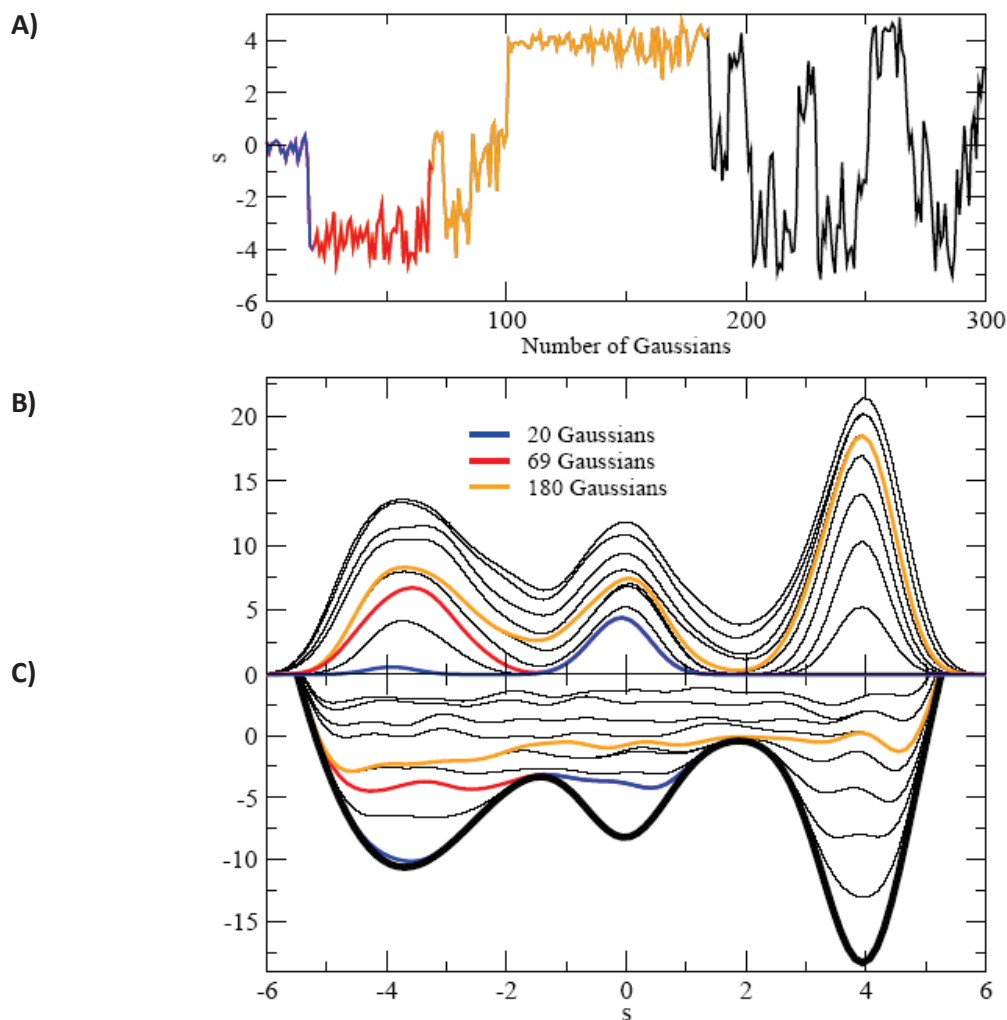


Figure 15. Principe de la métadynamique.

A) Visualisation de la trajectoire d'un système biaisé par un potentiel V évoluant à travers les 3 minima énergétiques (bleu, rouge et jaune) visibles sur le profil énergétique du système en C, en fonction du nombre de gaussiennes déposées. Avec s la valeur prise par la CV considérée, à l'instant t . B) Evolution au cours du temps du potentiel de biais V . La ligne bleue représente V lorsque le premier minimum est rempli, et que le système transite vers le second minimum. La ligne rouge représente V lorsque le second minimum est rempli. La ligne jaune représente V quand le profil énergétique est entièrement rempli et que le système devient diffusif. C) Evolution de la somme des potentiels V (lignes simples) et du potentiel externe (ligne épaisse), au cours du temps. Figure reproduite et modifiée de Laio et Gervasio 2008, avec l'autorisation de WIREs

Le choix des CVs est donc d'une importance capitale pour avoir une simulation fiable. Les CVs choisies dépendent du type de phénomènes que l'on veut étudier, du système dans lequel on se trouve, et elles peuvent être définies parmi tous les paramètres possibles du système. Le choix des CVs doit cependant satisfaire plusieurs critères :⁶³

- elles doivent permettre de distinguer nettement les états initiaux, finaux et intermédiaires
- elles doivent permettre de décrire les événements lents que l'on cherche à caractériser
- elles ne doivent pas être trop nombreuses, au risque de ralentir de manière importante la simulation, et de ne pas atteindre la convergence vers une moyenne stable

Dans le cas de phénomènes de reconnaissance moléculaire tel que la liaison d'un ligand dans le site actif d'une protéine, il est courant d'utiliser des CVs reliées à la géométrie du système : des distances, des angles, ou bien des angles dièdres.

Plus précisément, il est possible de choisir les CVs utilisées par les programmes de *docking* dans le cas de simulations de métadynamique : la distance entre le ligand et le site actif, et un ou plusieurs angles permettant de définir les orientations possibles du ligand à l'intérieur du site actif.⁶⁵ Si l'on cherche par contre à étudier des phénomènes plus complexes, dans des systèmes différents tels que des membranes, les CVs seront définies de manière complètement différente.⁶³

Plusieurs autres méthodologies ont été décrites telles que l'*umbrella sampling*,^{66; 67} ou la *steered MD* (SMD)⁶⁸. Dans le cas de la SMD, le ligand est traîné en dehors du site actif par l'action d'un potentiel de contrainte, alors que pour la méthode de l'*umbrella sampling* le chemin entre l'état libre et l'état lié est divisé en un nombre fini de plans qui sont échantillonnés indépendamment les uns des autres afin de reconstituer le chemin complet (Figure 16). Existe aussi la méthode de perturbation d'énergie libre qui a récemment montré des résultats probants pour l'identification d'imidazotriazines comme inhibiteurs de la SYK (*Spleen Tyrosine Kinase*) et l'optimisation de 17 de ces composés menant à des activités cellulaires de l'ordre du nanomolaire.⁶⁹

Le biais appliqué au système au cours de la simulation va aussi être utilisé afin de construire les cartes représentant l'énergie du système en fonction des différentes variables collectives choisies. Selon la précision recherchée, la valeur du biais peut alors être modifiée : une grande valeur pour obtenir des cartes énergétiques rapidement mais peu précises, et au contraire une faible valeur pour avoir une meilleure précision des états stables. De ce constat a émergé le protocole de « *well-tempered metadynamics* » dans lequel la valeur du biais est adaptée au cours de la simulation en fonction du chemin déjà parcouru.⁷⁰

La funnel-métadynamique, qui est un protocole récent de « *well-tempered metadynamics* », permet outre un échantillonnage accéléré, d'accéder à la valeur de ΔG_b° de manière quantitative.^{71,72}

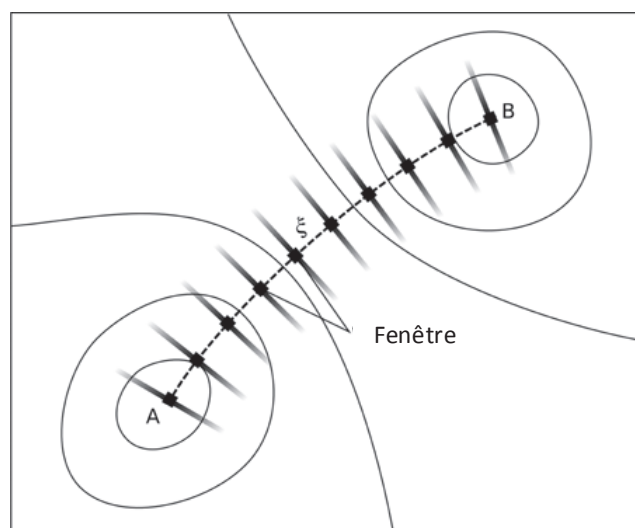


Figure 16. Représentation schématique de la dynamique moléculaire par *Umbrella Sampling*.

Séparation des coordonnées réactionnelles (ligne pointillée) entre deux états (minima A et B), par un nombre fini de « fenêtres ». Le système est échantillonné perpendiculairement aux coordonnées de la réaction. Figure reproduite de Johannes et al. avec l'autorisation de WIREs Comput Mol Sci.⁶⁶

Elle nous est apparue particulièrement adaptée pour simuler les phénomènes étudiables grâce aux observables RMN. De plus, elle permet d'obtenir la statistique des conformations de liaison et la cartographie de l'énergie au cours de la simulation. Je me concentrerai donc sur la présentation de cette méthode.

- La funnel-métadynamique

Grâce à la métadynamique, l'échantillonnage a pu être accéléré et il a été possible de cartographier l'énergie libre du système au cours de la simulation à partir du potentiel externe.⁶⁵ Cependant, seule une estimation qualitative de cette énergie a pu être obtenue car une fois le ligand dissocié, il explore le solvant et ne revient pas explorer le site d'interaction avant longtemps. Or si l'on souhaite obtenir une estimation quantitative de l'énergie libre de liaison d'un complexe PL, il est nécessaire d'échantillonner plusieurs événements d'association-dissociation afin d'obtenir une convergence de la moyenne de l'estimation des énergies libres.

Afin d'outrepasser cette limitation, V. Limongelli et ses collaborateurs ont présenté il y a 3 ans une nouvelle approche basée sur le concept de la métadynamique, appelée *funnel-métadynamique* (FM). Dans cette méthode, un potentiel extérieur appliqué au système se présente sous la forme d'un entonnoir, permettant de réduire l'espace à explorer lorsque le ligand se trouve sous sa forme non liée. Le potentiel est la combinaison d'une partie conique localisée au niveau du site actif de la protéine, et d'une partie cylindrique dirigée vers le solvant (Figure 17).⁷² Lorsque le ligand se situe dans l'intérieur de l'entonnoir le système n'est soumis à aucune contrainte, mais lorsque celui-ci

atteint les limites de l'entonnoir (les murs), alors un potentiel répulsif est appliqué au système, l'empêchant d'aller visiter les régions hors de l'entonnoir. Afin que le potentiel externe n'ait aucune influence sur le phénomène de liaison, il faut qu'il soit correctement dimensionné et positionné par rapport au site d'interaction. De plus, la diffusion de la protéine doit être inhibée (dans notre cas en fixant 3 atomes dans l'espace).

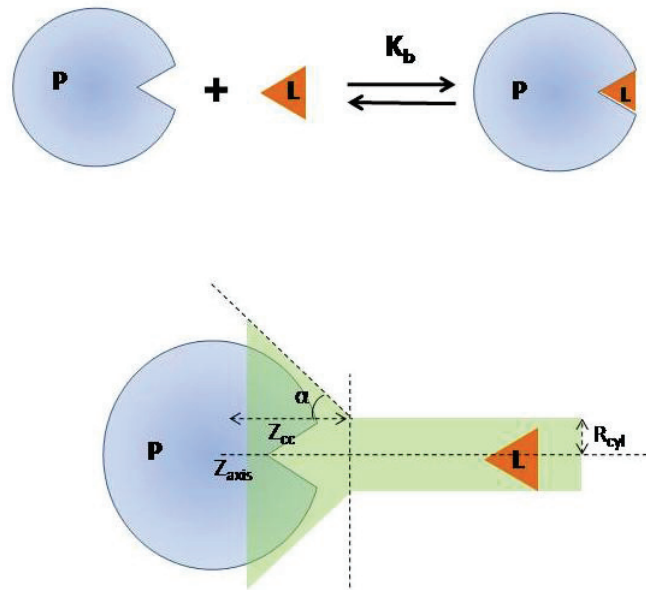


Figure 17. Représentation schématique du potentiel externe fixe utilisé en FM, adapté au phénomène de liaison protéine-ligand.

La forme du potentiel en entonnoir peut être adaptée en fonction du site récepteur étudié, selon les paramètres suivants : l'axe Z qui détermine l'orientation du potentiel, Z_{cc} la distance à laquelle le potentiel restreint passe d'une forme conique à une forme cylindrique, l'angle α définissant l'amplitude du cône, et R_{cyl} le rayon de la partie cylindrique. Figure reproduite depuis Limongelli et al. (2013) "Funnel metadynamics as accurate binding free-energy method." Proc. Natl. Acad. Sci. U.S.A., **110** (16):6358-6363.⁷²

Le potentiel permet d'assurer un nombre suffisamment important d'événements entre les différents états visités au cours de la simulation, conduisant à un profil énergétique quantitatif et bien caractérisé qui permet alors une estimation convergente de l'énergie libre de liaison ΔG_b^0 .

En FM, l'énergie libre absolue de liaison protéine-ligand est calculée de la manière suivante :

$$\Delta G_b^0 = \frac{-1}{\beta} \ln(C^0 K_b) \quad [24]$$

Ou K_b est la constante d'équilibre d'association du complexe, $C^0 = 1/1,660 \text{ \AA}^{-3}$ est la concentration standard correspondant à une concentration de 1 M, et $\beta = (k_B T)^{-1}$ avec k_B la constante de Boltzmann.

K_b peut aussi être exprimé en fonction du potentiel de force moyen (PMF), ici noté $W(z)$, de la manière suivante :

$$K_b = \frac{\int dz H_{site}(z) e^{-\beta W(z)}}{e^{-\beta W(z')}} \quad [25]$$

Où z' est la position de référence du ligand dans le solvant, $H_{site}(z)$ vaut 1 dans le site récepteur et 0 sinon. Pour ensuite exprimer K_b en fonction du PMF $W(z)$ à une dimension, il faut intégrer cette équation dans le plan xy comme décrit par Limongelli *et al.*, menant à l'équation suivante :⁷²

$$K_b = \pi R_{cyl}^2 \int_{z_{min}}^{z_{max}} dz e^{-\beta W(z)} \quad [26]$$

Avec πR_{cyl}^2 la surface du cylindre utilisé comme potentiel, qui tient compte du fait que dans le cas des simulations de métadynamique, l'utilisation du potentiel externe influe sur les paramètres liés à l'interaction protéine-ligand, proportionnellement à sa forme.

En termes d'énergie libre de liaison ΔG_b^0 , on obtient l'équation finale **27** :

$$\Delta G_b^0 = \Delta G - \frac{1}{\beta} \ln(\pi R_{cyl}^2 C^0) \quad [27]$$

Avec ΔG la différence d'énergie libre entre les états libre et lié, β et C^0 étant toujours des constantes.⁷² Il est donc à présent possible de calculer ΔG_b^0 .

Nous avons vu précédemment (p.20) que les expériences de RMN utilisant les CSP en fonction de la concentration totale en ligand permettaient de localiser les sites d'interaction au sein de la protéine. La FM apparaît donc ici particulièrement adaptée pour modéliser les interactions, les données RMN servant à positionner le potentiel externe fixe de la FM.

II. COMMENT DECOUVRE-T-ON DES MEDICAMENTS ?

Un objectif sous-jacent à notre étude est l'optimisation potentielle de l'affinité dans un contexte de conception et d'optimisation de petites molécules actives sur une protéine ou l'interprétation de l'affinité vis-à-vis de celle-ci. Comme la modélisation de l'affinité par MD est beaucoup plus réaliste que toutes les autres méthodes (*docking* rigide indus), on peut espérer en tirer des connaissances prédictives pertinentes pour baser une stratégie d'optimisation de ligand. Nos méthodes pourraient donc avoir un impact sur la chimie médicinale par exemple.

II. 1. Conception de molécules bioactives

Le développement complet d'un médicament depuis la source initiale est actuellement compris entre 12 et 15 ans et le coût approximatif serait de l'ordre de 1 milliard de dollars.⁷³ C'est donc un processus considéré comme long et risqué si l'on considère les taux d'échec même après les autorisations de mise sur le marché par les agences publiques. Nous allons définir ci-dessous les différentes approches possibles, ainsi que les étapes du procédé de conception de molécules à visée thérapeutique.

II. 1. 1 Les différentes approches existantes

Pour la découverte initiale de molécules actives, deux stratégies essentielles ont été adoptées.⁷⁴ Des criblages d'activité sont réalisés à partir de diverses prospectives dans des collections pouvant comprendre des extraits naturels (plantes, microorganismes, espèces marines) ou des chimiothèques de composés de synthèse obtenus diversement (par chimie combinatoire par exemple).

Un test d'activité peut comprendre des tests cellulaires *in cellulo* ou *in vivo* chez des organismes vivants (bactéries, plantes, insectes). L'expérimentation *in vivo* est la plus complexe à mettre en œuvre. Si les cellules ou les organismes présentent un phénotype particulier ou une pathologie spécifique, on parlera alors souvent de criblage phénotypique. Ces derniers ne renseignent pas nécessairement sur le ou les récepteurs moléculaires sur lesquels les molécules agissent.⁷⁵ Cependant, aux Etats-Unis par exemple, la connaissance des récepteurs moléculaires n'est indispensable auprès de la FDA (*Food and Drugs Administration*) que dans le cas des molécules qui agissent sur le système nerveux central.

Si par contre le ou les récepteurs ainsi que leur activité sont connus, on peut imaginer des tests d'activité *in vitro*. Ces tests permettent de détecter et éventuellement mesurer cette activité, en mettant en œuvre tout moyen physico-chimique. Les tests colorimétriques UV-Visible, la pH-métrie, ainsi que bien d'autres moyens analytiques peuvent être utilisés. Si celles-ci sont miniaturisables en microplaques à puits, elles peuvent être multipliées, automatisées et accélérées pour atteindre un très haut débit analytique.⁷⁶

II. 1. 2 Les étapes du processus de conception

La conception de molécules bioactives induit actuellement différentes disciplines allant de la chimie à la pharmacologie en passant par la biologie structurale, la chimie de synthèse et la modélisation moléculaire. La conception d'un médicament basée sur la connaissance structurale du récepteur biologique peut être décomposée en plusieurs étapes clés (Figure 18).⁷³

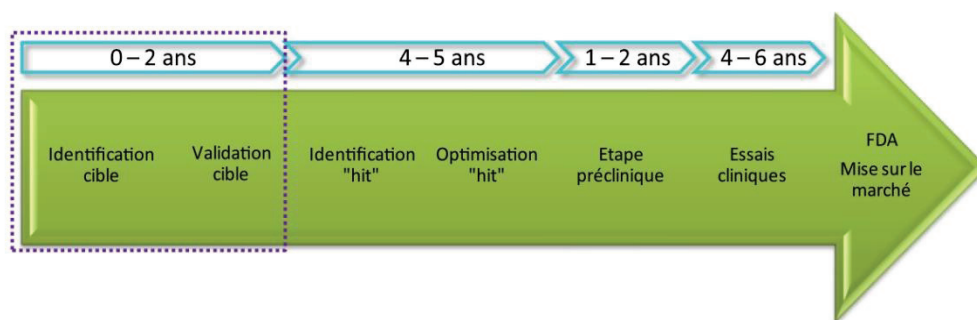


Figure 18. Schéma des étapes principales du procédé de conception d'une molécule bioactive.

Étapes de conception d'un médicament depuis sa source jusqu'à sa mise sur le marché dans le cas d'un criblage basé sur la structure du récepteur biologique. Dans le cas du criblage phénotypique, l'identification du hit se fait directement en observant ou non le phénotype responsable de la maladie, et cela sans nécessité d'identifier la cible.

Les molécules affines (*hits*) identifiées dont l'activité a été validée, doivent être optimisées afin d'obtenir la meilleure activité pour des doses minimales. Pour cela, différents analogues sont souvent testés pour comparer leurs activités dans un processus de type essai-erreur. Des compromis doivent à ce moment être anticipés pour s'adapter aux contraintes qui viendront des étapes précliniques. Lors de ces étapes, les meilleures molécules trouvées sont testées sur les modèles animaux pour déterminer leur efficacité et leur potentielle toxicité. Les essais cliniques effectués sur des modèles humains volontaires diront enfin si la molécule candidate est adéquate en terme de sécurité, d'effets secondaires, et le dosage optimal. Ce n'est qu'après validation de ces essais que les démarches pour l'obtention d'une autorisation de mise sur le marché peuvent commencer.⁷³

Sur l'ensemble de ces différentes étapes, nous nous intéresserons plus particulièrement aux étapes concernant l'identification et l'optimisation de molécules *hits* possédant un potentiel biologique pour la cible d'intérêt.^{73;77}

II. 2. Identification et optimisation de molécules affines (*hits*)

De manière générale, on distingue deux types de criblage: le criblage à haut débit de chimiothèques de type « *drug-like* » (HTS) et le criblage de chimiothèques de petites molécules appelées fragments (FBDD).

II. 2. 1 Le criblage haut débit

Le criblage HTS de grandes chimiothèques (de 250 000 jusqu'à 10^6 molécules) est encore aujourd'hui largement utilisé pour l'obtention de *hits* dont il faut par la suite confirmer l'activité réelle. En effet, les tests d'activité peuvent présenter des artefacts menant à de nombreux faux-positifs. Les molécules *drug-like* qui sont identifiées pour avoir une activité significative par criblage HTS sont donc par la suite étudiées de façon plus approfondie lors d'un criblage secondaire et leur action est caractérisée.⁷⁶ Statistiquement les molécules candidates sont souvent conformes à une règle connue sous le nom de « Règle des 5 ».⁷⁸ Cette règle établit que les molécules ayant un poids moléculaire ≤ 500 Da, un coefficient de partage octanol-eau (logP) inférieur à 5 et autres propriétés physico-chimiques résumées dans le Tableau 1, possèderaient des propriétés de solubilité et de biodisponibilité adaptées à un potentiel médicament.

II. 2. 2 La méthode des fragments (FBDD)

La méthode des fragments (FBDD pour *Fragment-Based Drug Design*), a été présentée comme une nouvelle alternative au criblage HTS. Elle repose sur le même principe de criblage, mais utilise des chimiothèques de molécules fragments (de 1 000 à 20 000 molécules) de plus faible poids moléculaire et structuralement moins complexes que les molécules *drug-like* présentes dans les chimiothèques utilisées pour le HTS.⁷⁹ En effet, la taille d'une molécule fragment ne dépasse pas 300 Da.⁸⁰ Par analogie avec la règle de Lipinski dans le cas du criblage HTS, une règle a été établie pour la méthode fragments, appelée « Règle des 3 », et qui suit les mêmes descripteurs chimiques réduits à 3 (Tableau 1).⁸¹

Cette méthode a l'avantage de donner un taux plus important de molécules *hits* (3 %) puisque les fragments sont des molécules plus simples avec peu de groupes fonctionnels, qui vont donc avoir

plus de chance de se lier mais probablement avec une plus faible affinité (Tableau 1).⁸² En effet si le criblage HTS permet l'identification à très faible fréquence de candidats ayant une affinité de l'ordre du nano ou micromolaire, le criblage de fragments permet l'identification de *hits* ayant une affinité comprise entre le micro et le millimolaire mais avec bien plus de possibilités de développement.

Tableau 1. Comparaison des méthodes HTS et FBDD.^{78; 81 80}

Méthode	HTS	FBDD
Taille maximale base de données	10 ⁶	2.10 ⁴
Taux de <i>hits</i>	≈ 0.1 %	≈ 3 %
Affinité de liaison	10 nm < K _D < 1 μM	10 μM < K _D < 10 mM
Propriétés physico-chimique	MW ≤ 500 g.mol ⁻¹	MW ≤ 300 g.mol ⁻¹
	Donneur H < 5	Donneur H < 3
	Accepteur H < 10	Accepteur H < 3
	logP < 5	logP < 3
	torsions < 5	torsions < 3

Avec logP le coefficient de partage octanol-eau, et P = C_{octanol}/C_{eau}

Page & Jencks ont formulé en 1971 l'hypothèse que lors de la formation d'un complexe, un fragment perd une quantité importante de son entropie (due aux rotations et translations), impliquant alors qu'il doit surmonter une barrière énergétique plus importante pour pouvoir se lier.^{83; 84} Par la suite, Jencks a pu démontrer que l'affinité d'un ligand peut être supérieure à la somme de l'affinité des fragments qui le composent.⁸³ Cette propriété de l'affinité a donné lieu aux premières expériences de structure-activité (SAR), décrites dans les années 90.⁸⁵

II. 2. 3 Les différentes méthodes analytiques pour le criblage de molécules

Nous avons vu en partie I.2.1 (p.15) qu'il existait de nombreuses techniques expérimentales permettant l'analyse de l'interaction PL. Ces techniques peuvent être utilisées dans le cas de criblage de molécules, selon les informations que l'on souhaite obtenir.

La RMN, est la méthode historique dans le FBDD. C'est par criblage RMN que les premiers résultats d'identification de fragments ont été obtenus dans le cas de la protéine FK506.⁸⁵ Par la suite, de nombreux résultats ont ensuite démontré son efficacité dans le cas de la méthode des fragments sur diverses protéines, mais aussi de manière plus générale dans le cas de la conception de médicaments basée sur la structure. L'équipe de S. W. Fesik, qui a majoritairement travaillé sur l'utilisation de la RMN pour la conception de médicaments et l'étude de complexe PL, a grandement contribué à son

développement par le biais de nombreuses publications dans ce domaine depuis les premiers résultats publiés en 1996.^{86; 87; 88} De plus, la RMN est une technique permettant de cribler des composés sur un large domaine d'affinité et, comme présenté précédemment, elle peut analyser les complexes PL en observant soit les signaux de la protéine, soit ceux du ligand.⁸⁹ Enfin, elle permet d'obtenir des informations sur le mode de liaison des composés criblés (partie I.1.2.2).²⁴

II. 2. 4 Optimisation des molécules *hits*

Le développement et l'optimisation des molécules *hits* passe notamment par l'adaptation de la stabilité du complexe PL, c'est-à-dire l'amélioration du temps de résidence moyen \bar{t}_R du ligand sur son récepteur.⁹⁰ Nous avons vu précédemment que ce \bar{t}_R est proportionnel à l'affinité d'un ligand (relations 6 p.9 et 9 p. 10). Cela passe aussi par l'amélioration des propriétés ADMET (Absorption Distribution Métabolisme Elimination Toxicité). La solubilité et la perméabilité (membranes, barrière intestinale, barrière hémato-encéphalique, etc) sont deux propriétés particulièrement cruciales dans le développement d'un médicament. La méthode utilisée pour l'optimisation des molécules identifiées par le criblage est basée sur les relations structure activité (SAR pour *Structure Activity Relationship*).⁷³ Les SAR consistent en général à déterminer une relation entre des motifs structuraux et des propriétés physico-chimiques ou une activité biologique significative. Des groupes de molécules possédant les mêmes propriétés sont définis afin de trouver le motif commun supposé responsable de cette propriété. De ces recherches s'initie souvent l'étape de développement de petites molécules fragments en molécules de type « *drug-like* ».

Pour développer une molécule *hit*, on peut ajouter des groupements fonctionnels afin de chercher à améliorer les propriétés ADMET citées précédemment, l'efficacité d'action et de liaison au récepteur.²⁴ On parlera dans ce cas d'évolution de fragment. Si plusieurs sites récepteurs ont été identifiés, il est aussi possible de lier plusieurs fragments qui ont été identifiés pour avoir deux sites proches l'un de l'autre sur le récepteur afin de gagner en affinité par l'ajout d'un « *linker* » entre ces molécules fragments comme utilisé dans le cas de la protéine FK506 (Figure 19).⁸⁵ On parlera alors plutôt de fusion de fragment.²⁴

- L'évolution de fragments

A la suite de l'identification d'un fragment se liant au site récepteur (*hit*), son développement par la stratégie d'évolution de fragments consiste à tester différentes substitutions sur ce *hit* (Figure 19) jusqu'à obtenir une affinité optimale. Le criblage associé au développement de ce *hit* se fera donc autour d'une même famille de molécules.⁷⁶

De cette méthode de développement de fragment, s'est dégagée la notion « d'efficacité ligand » (LE).^{79;91} La LE est un descripteur défini comme le rapport de l'affinité rapportée au nombre d'atomes lourds du ligand:

$$LE = \frac{\Delta G_b^0}{N} \quad [28]$$

Avec ΔG_b^0 l'énergie libre d'association du ligand pour la protéine et N le nombre d'atomes lourds du ligand. La taille des candidats médicaments est un descripteur important pour la biodisponibilité *in vivo*.⁹² En effet, les petites molécules diffusent en général mieux dans les organismes et les cellules que les grandes.

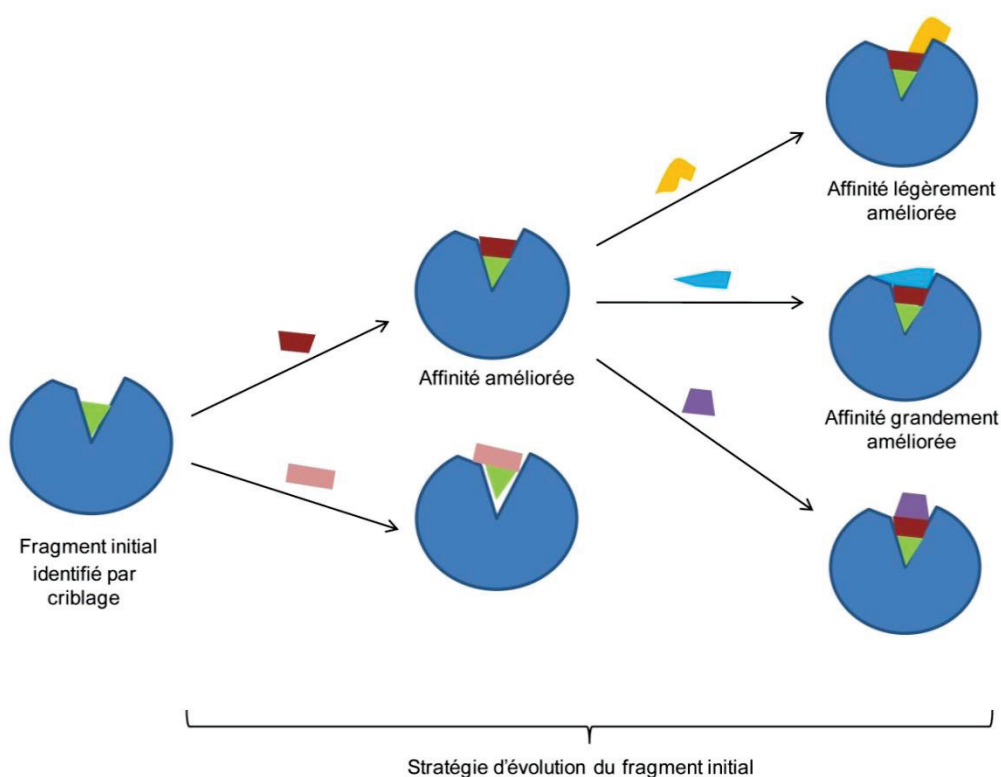


Figure 19. Stratégie d'évolution de fragment pour la conception d'une molécule *drug-like*.

Le fragment identifié affin pour le site récepteur est optimisé pas à pas par ajout de différents substituants afin d'améliorer son affinité pour le site récepteur, jusqu'à atteindre une molécule de type *drug-like*

- Liaison de fragments

Si plusieurs sites récepteurs ont pu être identifiés sur la cible, par RMN ¹⁵N-HSQC grâce à la mesure des CSPs, et que des fragments ont été identifiés comme interagissant dans ces sites adjacents, alors la stratégie de développement sera plutôt la liaison de fragments (Figure 20).⁸⁵ C'est-à-dire l'ajout d'un *linker* entre deux fragments identifiés affins pour 2 sites récepteurs adjacents de la

cible, afin de former une molécule de type *drug-like* possédant une affinité améliorée pour le récepteur. Cette technique est utilisée dès que possible afin de combiner les propriétés des différents fragments, ou leurs affinités sur différents (sous) sites de liaisons et pour augmenter l'affinité globale de la molécule grâce au principe de non-additivité de l'affinité.⁹³ Comme les fragments sont de faible poids moléculaire et ne possèdent qu'un nombre très restreint de groupement fonctionnels, il est plus aisé de trouver une relation entre un motif structural ou un groupe fonctionnel, et une propriété ADMET ou un degré élevé d'activité.⁷³

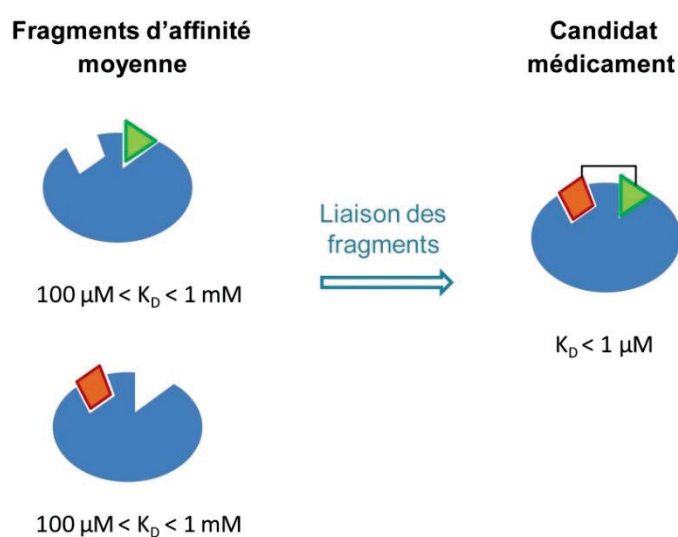


Figure 20. Schéma de la stratégie de liaison de fragments permettant la conception d'une molécule *drug-like*. Deux fragments d'affinité moyenne pour deux sites de liaison différents sont liés grâce à un *linker* pour la conception d'une molécule *drug-like* ayant une haute affinité pour la cible biologique.

Les techniques analytiques utilisées en SAR sont principalement la RMN et la cristallographie aux rayons X afin de caractériser les analogues étudiés.^{88; 93} Mais plus récemment, la dynamique moléculaire s'est révélée très utile dans ces étapes d'identification et d'optimisation de molécules affines, permettant une estimation théorique moins longue et moins coûteuse avec l'évolution des moyens de calcul.⁷⁷

II. 3. La dynamique moléculaire pour la conception de molécules affines

L'amélioration de la production et des coûts de développement est un des principaux défis industriels en général et aussi dans l'industrie pharmaceutique. En effet avec plus de 50 milliards de

dollars dépensés par an par l'industrie pharmaceutique, uniquement une vingtaine de médicaments voient le jour.⁹⁴ Les principales raisons de cette faible productivité sont diverses : connaissance trop limitée de la cible, manque d'activité ou de sélectivité de la molécule médicamenteuse, des propriétés inadaptées ou encore une toxicité imprévue.⁹⁵

L'utilisation de la dynamique moléculaire peut avoir de nombreux avantages dans la conception de molécules bioactives. Au-delà du criblage virtuel classique, les méthodes dynamiques permettent de prendre en compte le mouvement constant et la flexibilité de la protéine pour la prédiction directe des énergies d'association de petites molécules, mais aussi l'identification de possibles sites allostériques. La dynamique moléculaire permet de plus d'atteindre des détails sur la statistique et le mécanisme d'interaction à l'échelle microscopique qu'il ne serait possible d'atteindre avec des méthodes expérimentales.⁴²

Il est donc considéré que les méthodes numériques, en comblant le manque de données offertes par certaines techniques expérimentales macroscopiques, pourront probablement permettre d'améliorer la productivité dans la conception de molécules bioactives, surtout dans les étapes d'identification et d'optimisation en offrant une base beaucoup moins empirique.⁹⁵

II. 3. 1 Le criblage virtuel

L'association d'une molécule fragment à la surface de la protéine cible est gouvernée par des interactions non liantes (liaisons H, interactions électrostatiques et de van der Waals). Ces interactions doivent compenser les pénalités énergétiques dues à la désolvatation dans la zone d'interaction directe ligand-récepteur, ou à la perte de degrés de liberté pour que l'association soit possible. Nous avons vu en partie I.1.3 que certains concepts développés en bio-informatique permettent de modéliser les contraintes thermodynamiques et stériques pour une protéine cible de structure 3D connue. Grâce au *docking*, l'une des méthodes numériques les plus connues de criblage virtuel, les liaisons relatives de différents fragments pour une cible peuvent être prédites, permettant de classer les ligands par la valeur de leur énergie de liaison.⁷⁷ Cet outil est aussi utilisé comme outil de filtre pour les chimiothèques afin de réduire le nombre de molécules criblées en utilisant des descripteurs moléculaires pour prédire les propriétés ADMET des molécules criblées.⁷⁷

Actuellement, de nombreuses techniques de conception assistée par ordinateur sont utilisées dans le milieu industriel comme académique afin de réduire les coûts de production. Ces techniques de criblage virtuel peuvent être à la fois utiles dans les méthodes « *Structure-Based Drug Design* » (SBDD) et dans les méthodes « *Ligand-Based Drug Design* » (LBDD) (Figure 21).⁷⁷

- Le criblage virtuel en SBDD

L'approche SBDD est basée sur la connaissance de la structure tridimensionnelle de la macromolécule étudiée. Les ligands sont ainsi conçus et testés sur la base de leurs possibles interactions avec des sites récepteurs. La structure de base peut être déterminée par cristallographie aux rayons X ou RMN par exemple, ou bien prédite par des méthodes numériques (méthode d'homologie).

Cette approche comporte les méthodes numériques de criblage virtuel, et la conception de ligand « de novo » (Figure 21). Les méthodes de criblage virtuel classique, tel que le *docking*, sont très utilisées dans l'étape d'identification de molécules *hits* ayant une activité biologique significative. Une chimiothèque est criblée virtuellement contre la structure de la macromolécule d'intérêt, et les molécules sont ensuite classées selon l'estimation de leur énergie de liaison calculée par une fonction de *scoring* empirique, qualitative.

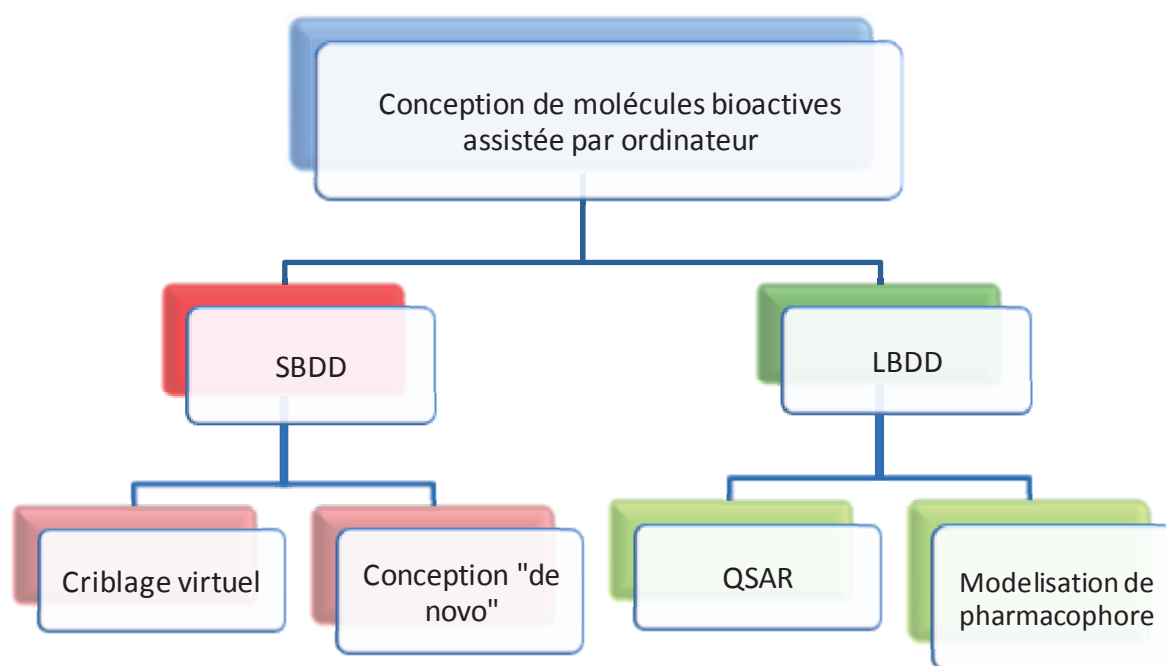


Figure 21. Représentation schématique des différentes approches numériques pour la conception de molécules bioactives.

2 approches principales basées soit sur la connaissance de la structure du récepteur biologique, soit sur les ligands se liant à la protéine d'intérêt. En *Structure-Based Drug Design* (SBDD) deux méthodes sont possibles : le criblage virtuel de chimiothèques de molécules, ou la conception de ligand « de novo ». En *Ligand-Based Drug Design* (LBDD) il est soit possible d'étudier les pharmacophores pour orienter la conception, soit la méthode Quantitative Structure-Activity Relationship (QSAR) pour trouver une relation empirique entre des motifs structuraux et des propriétés pharmacologiques spécifiques.

Cette fonction de *scoring* prend généralement en compte les termes d'interactions de vdW et les potentiels de Coulomb qui sont pondérés de manière empirique.⁹⁶ Cette méthode permet aussi l'estimation des conformations de liaison à l'intérieur du site actif du récepteur. Cependant la dynamique du squelette de la protéine n'est pas prise en compte en dehors des chaînes latérales des acides aminés, et les ligands sont « dockés » sur une protéine au squelette rigide.⁷² C'est une méthode largement utilisée, de manière complémentaire au criblage HTS expérimental.⁹⁷

Pour l'optimisation des fragments ou la conception de ligands spécifiques, sont plutôt utilisées les méthodes de conception « de novo ». Ces méthodes s'appuient sur la connaissance des acides aminés et de leur orientation à l'intérieur du site actif pour concevoir ou optimiser un fragment qui corresponde au mieux à ce site actif en fonction des interactions pouvant être mises en place entre les deux partenaires biologiques. Les outils numériques sont utilisés pour l'analyse détaillée du site récepteur, ou bien l'évaluation des affinités de ligands spécifiques selon des critères particuliers (forme, complémentarité électrostatique, etc).^{42;77;97}

- Les méthodes numériques en LBDD

A l'inverse, le LBDD est une approche utilisée lorsque la structure tridimensionnelle du récepteur est inconnue et qui se base sur la connaissance des molécules connues comme se liant à la cible d'intérêt.⁹⁷ Le but est de définir un modèle de criblage fiable en étudiant un panel large de composés possédant des motifs structuraux similaires liés à leur activité biologique. Dans cette approche, il est donc admis que des composés de structure proche, aurait une activité biologique et une interaction similaire avec la cible.⁷⁷ Lorsque l'on se réfère à ce type d'approche, les techniques numériques les plus largement utilisées sont les méthodes basées sur les pharmacophores (portion de la molécule active responsable de l'action physiologique), et la méthode QSAR (*Quantitative Structure-Activity Relationship*) dérivée de la SAR où une série de molécules de différentes structures sont étudiées afin de chercher une relation empirique entre des descripteurs structuraux et des propriétés spécifiques.^{77;97}

Au cours des dernières décennies les approches SBDD et LBDD ont évolué séparément. Cependant leur utilisation combinée pourrait sans doute permettre d'avoir un criblage plus efficace menant à de meilleures affinités pour la cible.⁹⁵

II. 3. 2 Nouvelles méthodes basées sur la dynamique moléculaire

L'utilisation de la MD pour la prédiction des meilleurs ligands est une nouvelle forme de criblage virtuel. Les principales limitations des méthodes basées sur la MD sont dues à la puissance de calcul nécessaire pour effectuer des simulations de l'ordre de 1 μ s ou plus.⁴² Avec l'émergence de

nouveaux systèmes de calcul comme Anton,⁵⁷ ARCHER,⁹⁸ ou SunwayTaihuLight (le supercalculateur le plus puissant au monde en juin 2016)⁹⁹, ainsi que l'usage de nouveaux processeurs graphiques (GPU), les méthodes de MD ont évolué afin de pouvoir simuler et visualiser des événements sur une plus grande échelle de temps. Actuellement, les méthodes de dynamique moléculaire présentées en première partie d'introduction, telles que la dynamique moléculaire accélérée ou la métadynamique permettent d'accéder à des mécanismes biologiques dits « longs » de l'ordre de la milliseconde et plus. De tels phénomènes incluent le repliement d'une chaîne protéique par exemple, ou bien les échanges conformationnels ayant lieu lors de l'association avec une petite molécule. Des contraintes (ou biais) permettent d'accélérer l'évolution du système.⁵⁹

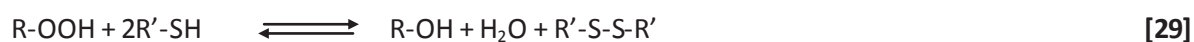
De plus, la MD permet d'obtenir des informations viables sur le(s) chemin(s) d'association/dissociation des ligands puisque dans ce cas les atomes de la protéine sont mobiles, et non pas fixes dans l'espace comme en *docking*.¹⁰⁰ La dynamique interne de la protéine étant conservée il est possible de déterminer les différentes conformations d'association possibles à l'intérieur du site récepteur en mouvement, et de visualiser les possibles modifications de conformation induites par la liaison d'une molécule ligand.¹⁰⁰ Il est aussi possible d'identifier d'éventuels états de transition ainsi que leurs énergies, et ainsi reconstruire de manière réaliste le chemin d'association.¹⁰¹ La MD peut-être utilisée comme méthode de criblage, mais peut aussi venir en complément d'un criblage primaire effectué par *docking* pour la validation des *hits* identifiés, ou leur optimisation.¹⁰⁰

III. MODELE DE L'ETUDE : LA PEROXYREDOXINE5 HUMAINE

Notre projet va être plus particulièrement focalisé sur l'étude d'interactions avec l'isoforme 5 de la peroxyrédoxine humaine, hPRDX5. Les peroxyrédoxines ont été décrites comme ayant un rôle médiateur dans l'inflammation post-ischémique cérébrale. Au sein du laboratoire, son protocole de production était bien maîtrisé ainsi que son marquage isotopique ^{15}N et/ou ^{13}C . Un criblage par RMN d'une collection de fragments avait été effectué au laboratoire en 2010. Ce criblage a donné 9 fragments *hits* dont 6 ont été capables d'induire des perturbations de déplacement chimique sur les protons amides de la protéine.⁹¹ De cette étude s'est dégagé un ligand particulier: le *o*-pyrocatechol (CAT), qui induisait les plus grandes CSPs visibles en RMN, et qui nous servira de base pour l'étude proposée ainsi que la recherche et l'optimisation de nouveaux ligands.

III. 1. Les peroxyrédoxines

Les peroxyrédoxines (PRDX) sont des enzymes (peroxydases) antioxydantes qui permettent la décomposition du peroxyde d'hydrogène (H_2O_2) en eau, ainsi que différents hydroperoxydes organiques (ROOH) et peroxynitrites (RNOO^\cdot) selon la réaction suivante :



Elles ont pour rôle de protéger les cellules contre les espèces réactives oxygénées (ROS) et les espèces réactives azotées (RNS) créées par le stress oxydant.^{102; 103} Elles réduisent plus de 90% des peroxydes cellulaires, avec des constantes catalytiques de l'ordre de $10^7 \text{M}^{-1} \cdot \text{s}^{-1}$.^{104; 105; 106} Les PRDX peuvent avoir d'autres rôles biologiques : chaperones, partenaires de liaison, activateurs d'enzyme ou régulateurs redox, du fait de leurs différents états d'oligomérisation que nous évoquerons par la suite.^{107; 108}

III. 1. 1 Classification des peroxyrédoxines

- Les 6 familles de peroxyrédoxines humaines

Les PRDX ont été identifiées dans des cellules de levures, de plantes, de bactéries et d'archées, ainsi que chez l'animal et l'Homme. Certains de ces organismes possèdent plusieurs isoformes de ces protéines, et chez les mammifères, on trouve 6 sous familles de PRDX de gènes distincts.¹⁰⁹ Les 6 isoformes de peroxyrédoxines humaines (hPRDX) sont présentes dans de nombreux compartiments cellulaires. Les peroxyrédoxines hPRDX1, hPRDX2 et hPRDX6 sont majoritairement

localisées dans le cytosol. hPRX5, que l'on retrouve aussi dans le cytosol, va aussi être présente dans les mitochondries et le peroxysome. hPRDX3, elle, est localisée uniquement dans la mitochondrie, et hPRDX4 dans le réticulum endoplasmique.¹⁰⁸ Les formes les plus exprimées dans les tissus biologiques de tout type sont hPRDX1 et hPRDX3.

Séquences

Le site actif des PRDX est hautement conservé.¹¹⁰ En termes de séquence, on retrouve un motif d'acides aminés particulier au niveau du site actif qui est conservé : **PXXXTXXCp** (Figure 22).^{103;}

¹¹¹ Ce motif contient un résidu Pro et un résidu Thr. La conservation de ce motif permet la conservation du réseau d'interactions de vdW impliqué dans le site actif, mettant en jeu la cystéine peroxydatique (Cp) avec les acides aminés Pro40, Thr44, Cys47 et Arg127 (numérotation de la hPRDX5) (Figure 23).¹¹² Thr44 et Pro40 sont toutes deux présentes dans la dernière boucle de l'hélice α dans laquelle est localisée la Cp. Chacun de ces acides aminés a un rôle bien particulier. Pro40 par exemple, de par son caractère hydrophobe venant de son cycle pyrrolidine, permet de limiter l'accessibilité de la Cp au solvant, dont seul le thiolate est accessible, comme visible en Figure 23.

Le résidu Arg127 quant à lui, va permettre d'abaisser le pKa de la Cp du fait du caractère hydrophile et basique de son groupe guanidinium, et donc stabiliser la charge négative de l'atome de soufre de la cystéine catalytique.¹⁰⁹ Thr44 et Arg127, de par leur proximité avec l'atome d'oxygène de H_2O_2 lors de sa liaison avec hPRDX5, peuvent interagir avec cet oxygène et ainsi aider à la stabilisation de l'état de transition du complexe protéine-ligand.¹¹³ D'autres résidus localisés en dehors du site actif sont aussi conservés tels que Glu60, Ser74 et Trp84. Ser74 et Trp84 sont localisés au niveau de l'interface dimérique, et ont été proposés comme participant au réseau de liaisons H formé entre la Cp et les résidus conservés à l'intérieur du site actif (Figure 23). Ils sont donc indirectement impliqués dans l'interaction PL.¹¹¹

Structures

Les membres de la famille des hPRDX partagent une base structurale commune de type thioredoxine (Trx) ou encore dit « repliement de Rossmann » avec un résidu cystéine dite peroxydatique (Cp) invariable.¹¹⁴ Cette Cp est la cystéine portant le groupement thiol (ou thiolate en l'occurrence) réagissant directement avec le peroxyde. Leur structure tertiaire est largement conservée, mais peut différer au niveau de boudes et des extrémités N-terminale (*N-term*) et C-terminale (*C-term*). En effet, la structure centrale des hPRDX comprend 7 brins β et 5 hélices α qui sont agencés de manière similaire. En repliement Rossmann, un feuillet β antiparallèle est formé de 5 brins ($\beta 5$ - $\beta 4$ - $\beta 3$ - $\beta 6$ - $\beta 7$), dont une face est recouverte par $\beta 1$ - $\beta 2$ - $\alpha 1$ et $\alpha 4$ et l'autre face par $\alpha 2$, $\alpha 3$ et $\alpha 5$ (Figure 24.A).

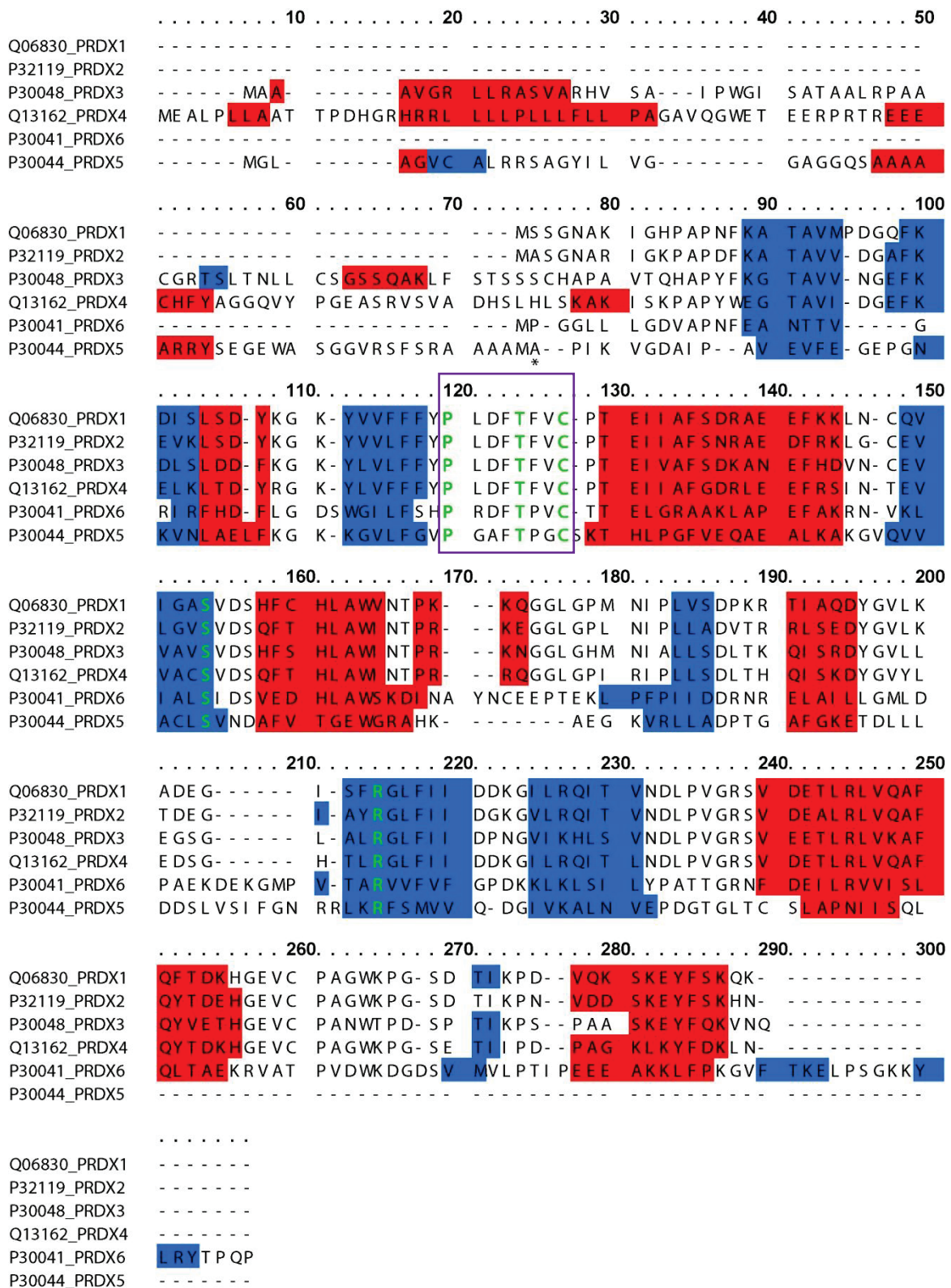


Figure 22. Alignement de séquence des 6 isoformes de hPRDX et visualisation des acides aminés conservés.

Séquences données avec leur fragment d'adressage mitochondrial, prises sur UniProt (codes donnés en début de séquence). Début de la séquence de hPRDX5 utilisée notifiée d'un *. Visualisation des éléments de structure secondaire des 6 isoformes de hPRDX : les feuillets β en bleu et les hélices α en rouge. En vert sont représentés les acides aminés conservés dans les isoformes de hPRDX. Dans la région du site actif on trouve les acides aminés Pro40, Thr44, Cys47 et Arg127, et dans le reste de la séquence, Ser74 et Trp84 (numérotés par rapport à leur place dans la séquence de hPRDX5) qui participent au réseau de liaisons H intervenant dans le site actif. L'alignement de séquences a été effectué à l'aide du site PRALINE en utilisant la méthode PSI-BLAST.

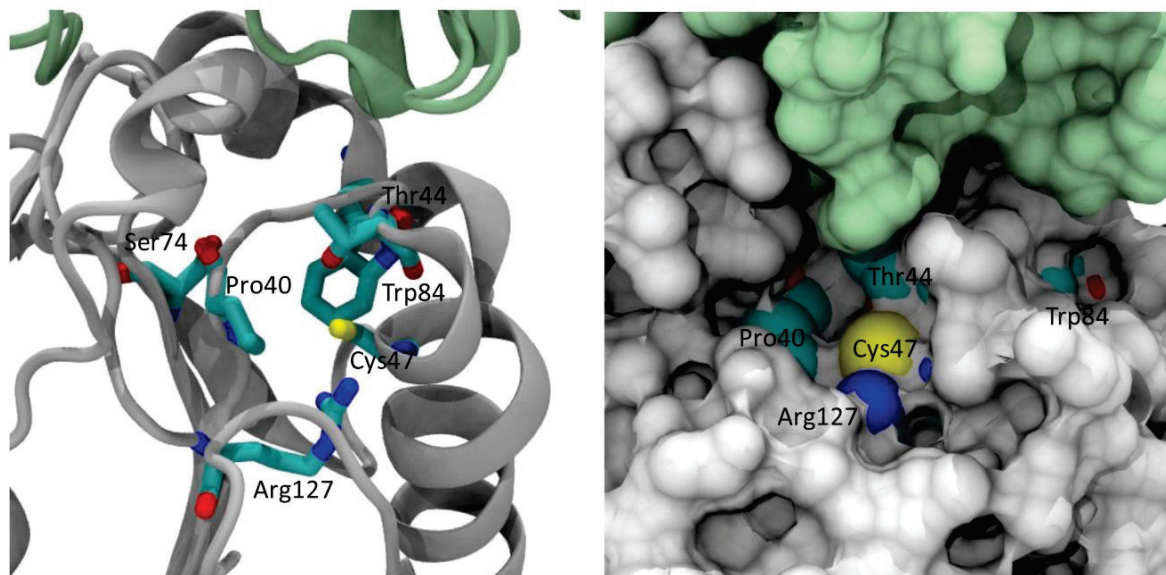


Figure 23. Représentations des résidus conservés dans les hPRDX sur le modèle structural de hPRDX5 et visualisation de l'accessibilité de la Cp.

Visualisation du site actif de hPRDX5 sous forme cartoon (gauche) et surface (droite). Les acides aminés conservés sont représentés en bâtons (gauche) et sphères de vdW (droite). Les deux chaînes de l'homodimère sont colorées en blanc (A) et vert (B).

Etant donné que les brins $\beta 5$ et $\beta 1$ interagissent ensemble, on parle plus communément d'un feuillet à 7-brins. Lorsque les PRDX sont natives, leur cystéine peroxydatique est localisée au niveau du début de l'hélice $\alpha 2$ (Figure 24.A).¹¹² De manière plus spécifique, c'est toute la géométrie du site actif centré autour de cette Cp qui est conservé (Figure 23).¹⁰⁹

Si l'on s'intéresse à présent à leur structure quaternaire, on peut mettre en évidence de plus grandes différences.¹⁰⁷ Depuis près de 20 ans, près de 40 structures cristallographiques de PRDX ont été publiées, dans différents états d'oxydation allant du thiol libre (Cp-S⁻) à l'acide sulfonique (Cp-SO₃H), et dont les états d'oligomérisation varient du monomère au décamère, voire même au dodécamère.^{115;116}

L'oligomérisation des PRDX peut se faire selon deux interfaces différentes que nous nommerons de type A et de type B (Figure 24).^{109;112;117} L'interface dimérique de type A se caractérise par la mise en contact de l'hélice $\alpha 3$ de l'un des monomères avec l'hélice $\alpha 5$ de l'autre monomère, ce qui va aussi impliquer les résidus appartenant aux boudes $\beta 3$ - $\alpha 2$ et $\beta 5$ - $\alpha 4$. Ce type d'interface est observé dans le cas de hPRDX5 et dans les cas d'oligomérisation mettant en jeu les deux types d'interfaces. Dans l'interface dimérique de type B, il y a association antiparallèle des brins $\beta 7$ des deux monomères en *C-term* conduisant à la formation d'un feuillet β à 10 brins. Ce type d'oligomérisation est observé pour les peroxyrédoxines hPRDX1, hPRDX2, hPRDX3, hPRDX4 et hPRDX6, mais n'a jamais été observé pour hPRDX5 qui possède en *C-term* une boucle qui empêche la mise en contact des deux brins β . Le degré d'oligomérisation dépend de plusieurs facteurs physico-chimique *in vitro* comme l'état redox, la force ionique, la concentration, le pH,...

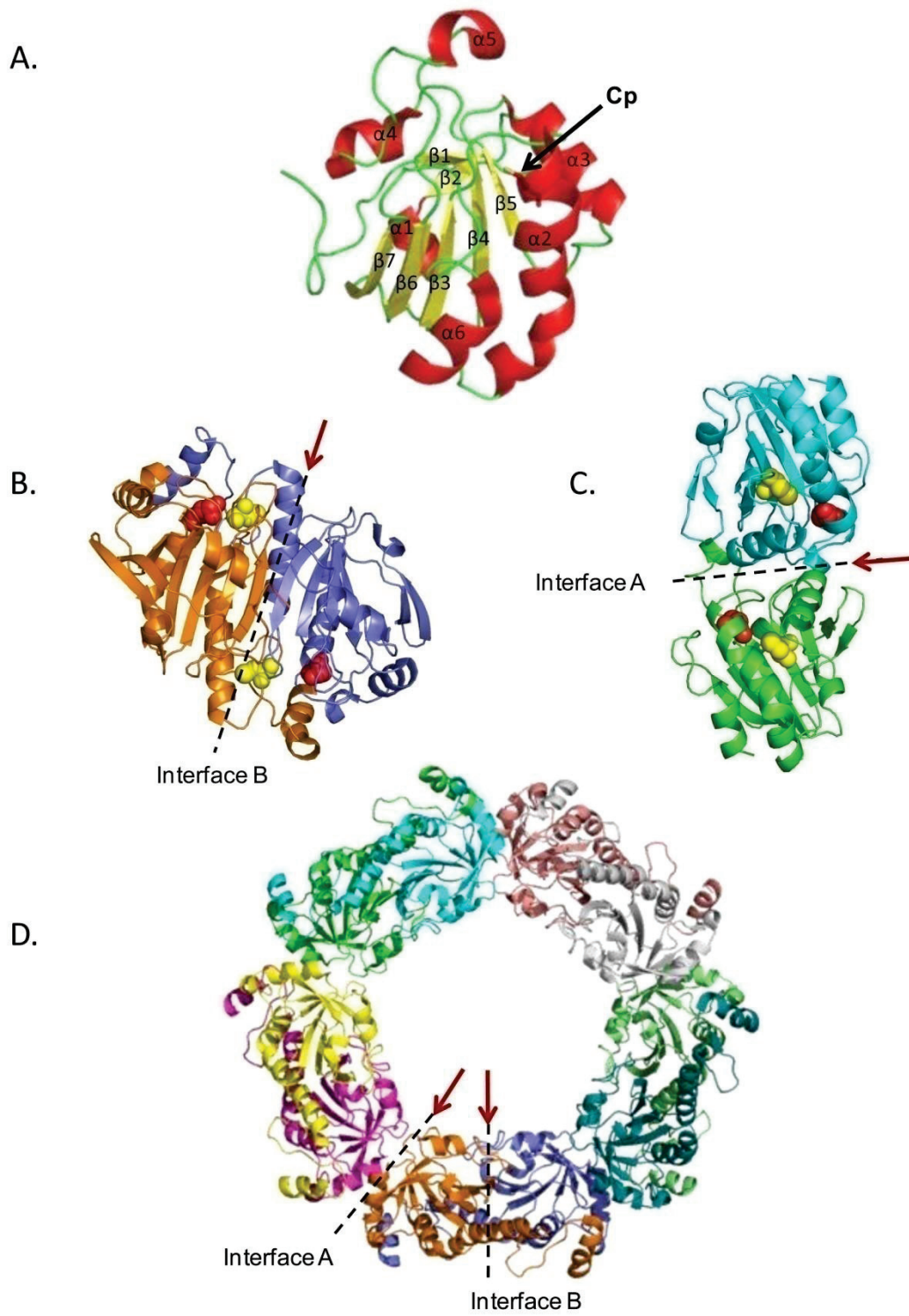


Figure 24. Structures des hPRDX et de leurs interfaces dans les structures quaternaires.

Les cystéines catalytiques (Cp) sont représentées sous forme de sphères rouge et les cystéines de régénération (Cr) en jaune. A) Représentation de l'une des deux chaînes de l'homodimère symétrique de hPRDX5 (3MNG.pdb) avec légende des structures secondaires. B) Représentation de hPRDX2 sous forme dimérique, avec visualisation de son interface de type B. C) Homodimère symétrique de hPRDX5 avec visualisation de son interface dimérique de type A. D) hPRDX2 décamerique avec visualisation des deux types d'interfaces A (bleu) et B (vert) présentes dans l'oligomère.

Par exemple, dans le cas des peroxyrédoxines 1 à 4, leur oligomérisation en décamères ou dodécamères peuvent mener à des assemblages de haut poids moléculaire comme le dodécaèdre par exemple (assemblage de 12 décamères) observé dans le cas de la hPRDX2 dans les érythrocytes humains en présence d'ATP et d'ions Mg^{2+} par microscopie électronique.^{107;118}

- Les 3 classes de peroxyrédoxines

On peut classer ces 6 familles de hPRDX selon 3 classes qui partagent le même mécanisme catalytique de base, dans lequel la Cp du site actif est oxydée par le substrat pour former un acide sulfénique Cp-SOH.¹¹⁹ Les 3 classes vont se distinguer par la manière dont va être réduit l'acide sulfénique Cp-SOH en thiol Cp-SH pour régénérer le résidu cystéine initial, et donc le nombre et la position des résidus cystéine qui participent au mécanisme catalytique (Figure 25).^{109;120;121}

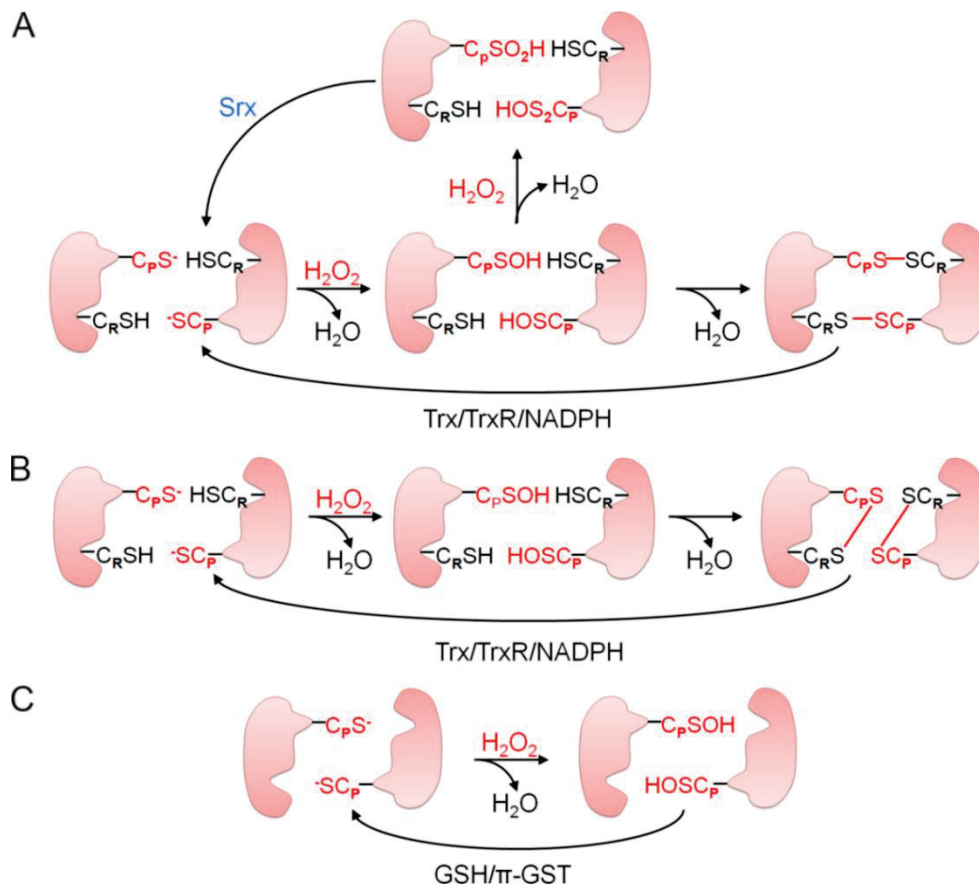


Figure 25. Mécanisme catalytique des PRDX en fonction de leur classe.¹

2-Cys PRDX (A), 2-Cys atypiques (B) et 1-Cys (C). La Trx agit à travers la TrxR et le NADPH.

Trx : Thioredoxine ; TrxR : Thioredoxine réductase ; GSH : Glutathion ; π -GST: isoforme Pi de la glutathione S transferase ; NADPH : Nicotinamide Adénine Dinucléotide Phosphate réduit

¹ This research was originally published in the Journal of Biological Chemistry by Rhee et al. Peroxiredoxin Functions as a Peroxidase and a Regulator and Sensor of Local Peroxides Journal of Biological Chemistry. 287, 4403-4410, 2012 © the American Society for Biochemistry and Molecular Biology

Dans la classe des PRDX à 2 cystéines (dites typiques) on retrouve les hPRDX 1 à 4 (Figure 25.A), qui possèdent 2 résidus cystéines conservés : une cystéine peroxydatique (Cp) localisée en *N-term* une cystéine de régénération (Cr) localisée en *C-term* de l'autre sous unité. Ces deux résidus cystéine vont tous deux participer à l'activité catalytique de l'enzyme. Cp-SH est tout d'abord sélectivement oxydée par H₂O₂ en acide sulfénique Cp-SOH, puis Cr-SH va venir former un pont disulfure intermoléculaire avec Cp-SOH afin d'être réduit par la thiorédoxine (Trx) associée à la thiorédoxine réductase (TrxR) et le NADPH dans un mécanisme que l'on définira par la suite. On peut noter que pour cette classe de PRDX, il leur est possible de réagir avec une autre molécule de H₂O₂ et ainsi être sur-oxydée sous forme d'acide sulfinique Cp-SO₂H. Dans ce cas, l'enzyme se trouve inactive et nécessite l'action de la sulfirédoxine (Srx) pour se régénérer et retrouver son activité peroxydase.¹¹⁵

hPRDX5 appartient à la classe des PRDX à 2 cystéines dites atypiques (Figure 25.B)¹²²: ici, seule la cystéine conservée en *N-term* (Cp) est présente, mais l'enzyme aura besoin de la Cr située dans la même chaîne polypeptidique afin de pouvoir activer sa fonction catalytique. Au contraire des PRDX de la première classe, les deux cystéines Cp et Cr sont localisées dans la même sous-unité et ne vont pas former des ponts disulfures intermoléculaires, mais des ponts intramoléculaires. De la même manière que pour les PRDX à 2 cystéines, le système Trx/TrxR/NADPH va ensuite permettre sa réduction.

Enfin, dans la dernière classe, dite des PRDX à 1 cystéine, on retrouve hPRDX6 (Figure 25.C). Dans cette classe de PRDX, seule la cystéine peroxydatique est conservée en *N-term*, et ne nécessite pas l'intervention de cystéine régénérative pour son activité catalytique. Dans son cas, la Cp oxydée ne formera pas de pont disulfure, et sera réduite directement par une molécule donneuse d'électrons possédant un groupement thiol, de type glutathion GSH.¹⁰⁹

III. 1. 2 Activité des peroxyrédoxines

- Les différentes fonctions biologiques des peroxyrédoxines

Comme évoqué précédemment, les PRDX sont une famille de protéines qui éliminent et donc détoxifient le milieu de H₂O₂, des peroxy-nitrites,¹²³ et des hydroperoxydes de nature organique. Ces enzymes qui peuvent être qualifiées d'antioxydantes protègent donc les cellules du stress dit oxydant dû à la production de ROS et RNS en réduisant plus de 90% des peroxydes cellulaires.¹⁰³ De plus leur large distribution, leur haut niveau d'expression chez *Escherichia coli* et leurs faibles valeurs de constantes de Michaelis rend leur rôle dans la défense antioxydante et la signalisation redox d'autant plus important.^{124; 125; 106} Elles vont donc avoir un fort potentiel thérapeutique dans les cas de maladies cardiovasculaires ou de cancers impliquant des ROS.¹²⁶ En effet, en inhibant les PRDX on devrait pouvoir restaurer la fonction apoptotique des cellules en les sensibilisant aux ROS. Mais les

PRDX n'ont pas seulement des fonctions biologiques liées à leur activité de peroxydase, et en réalité, diverses fonctions biologiques ont été publiées concernant les PRDX en fonction de leur localisation, de leur état redox et de leur degré d'oligomérisation.^{108; 121} Elles ont par exemple un rôle dans la modulation de l'expression de cytokines et dans la transduction du signal intracellulaire.¹²⁶ En régulant le taux de peroxydes dans la cellule elles vont avoir une action sur la cascade de signalisation conduisant à la prolifération cellulaire, la différenciation des cellules et leur apoptose.¹⁰⁹

- Dépendance des peroxyrédoxines aux thiorédoxines

Les PRDX ne sont pas catalytiquement actives sous leur forme oxydée disulfure. Elles doivent alors être régénérées par un réducteur adéquat comme une Trx après réaction avec le peroxyde d'hydrogène H₂O₂ selon les deux réactions suivantes:¹⁰⁸



La Trx est une protéine ubiquitaire de 12 kDa qui va réduire les PRDX par réduction des ponts disulfures. Elle possède un site actif caractérisé par la présence de 2 résidus cystéine qui se font face et qui peuvent alors s'oxyder et former un pont disulfure par oxydation des deux groupes fonctionnels thiols (Figure 26). Au cours de cette réaction, un autre pont disulfure est réduit.

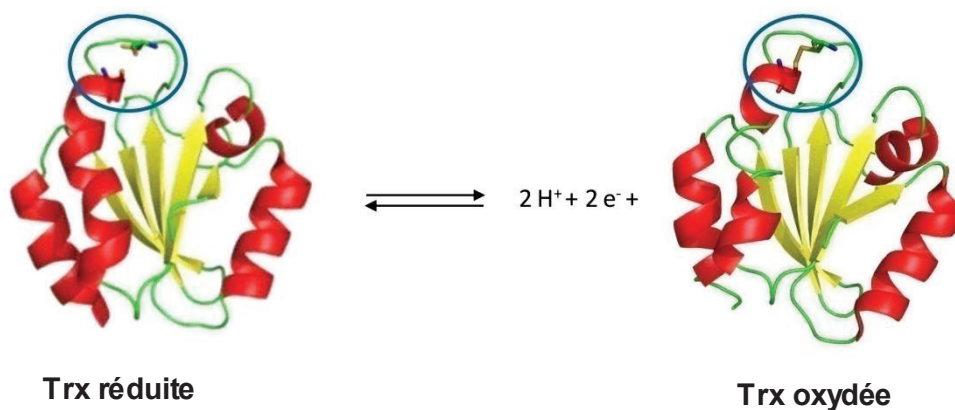


Figure 26. Structures oxydée et réduite de la Trx humaine, en équilibre d'oxydoréduction.

Structures de la thiorédoxine humaine réduite Trx_{red} (3TRX.pdb) et oxydée Trx_{ox} (1TRS.pdb). Visualisation du motif catalytique Cp-X-X-Cp entouré et la formation du pont disulfure lors de l'oxydation.

La Trx_{ox} est ensuite régénérée en Trx_{red} par l'action de la TrxR qui fonctionne en deux temps avec le NADPH (+H⁺) via un coenzyme (le FAD), comme suit (Figure 27)¹²⁸:

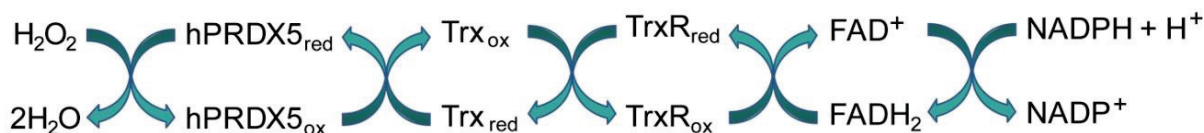
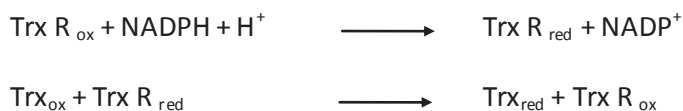


Figure 27. Schéma récapitulatif du cycle catalytique de régénération des hPRDX 1 à 5 par la Trx et la TrxR au travers du NADPH.

- Cycle catalytique

Le cycle catalytique complet des PRDX comprend de 3 à 5 étapes, selon que l'on considère ou non une suroxydation, et 2 états conformationnels distincts des PRDX. Ces deux états sont l'état dit FF (*Fully-Folded*) et l'état LU (*Locally Unfolded*) (Figure 28).

Les 3 étapes de base du mécanisme catalytique des PRDX sont les étapes de peroxydation, de résolution et de recyclage. Lorsqu'une PRDX est suroxydée (c'est à dire que l'acide sulfénique R-SOH est oxydé en acide sulfinique R-SO₂H), une étape dite de « résurrection » est nécessaire impliquant une enzyme supplémentaire, la sulfiredoxine (Srx), avant de pouvoir reprendre part au cycle catalytique classique.¹²⁹

La première étape du cycle catalytique, appelée peroxydation, correspond à l'attaque nucléophile de la Cp sous forme Cp-S⁻, sur le substrat (H₂O₂). Cp est alors oxydée en Cp-SOH et il y a formation d'eau H₂O (alcool dans les cas des peroxydes organiques). Cette étape est uniquement possible lorsque la protéine est complètement repliée, induisant alors un abaissement du pKa de la Cp (normalement de 8.3) qui se trouve essentiellement sous forme thiolate à pH 7.

Durant l'étape de résolution, la protéine se trouve dans une conformation localement dépliée au niveau de la fin de l'hélice portant Cp. La Cp est ensuite attaquée par un deuxième groupement thiol, créant ainsi un pont disulfure (intra ou intermoléculaire). L'eau (ou l'alcool) formée à l'étape précédente est alors éliminée. Dans le cas des PRDX à 2 cystéines, ce groupement thiol est celui de la Cr, tandis que pour les PRDX à 1 cystéine (PRDX 6), il provient d'une petite molécule soufrée de type glutathion.

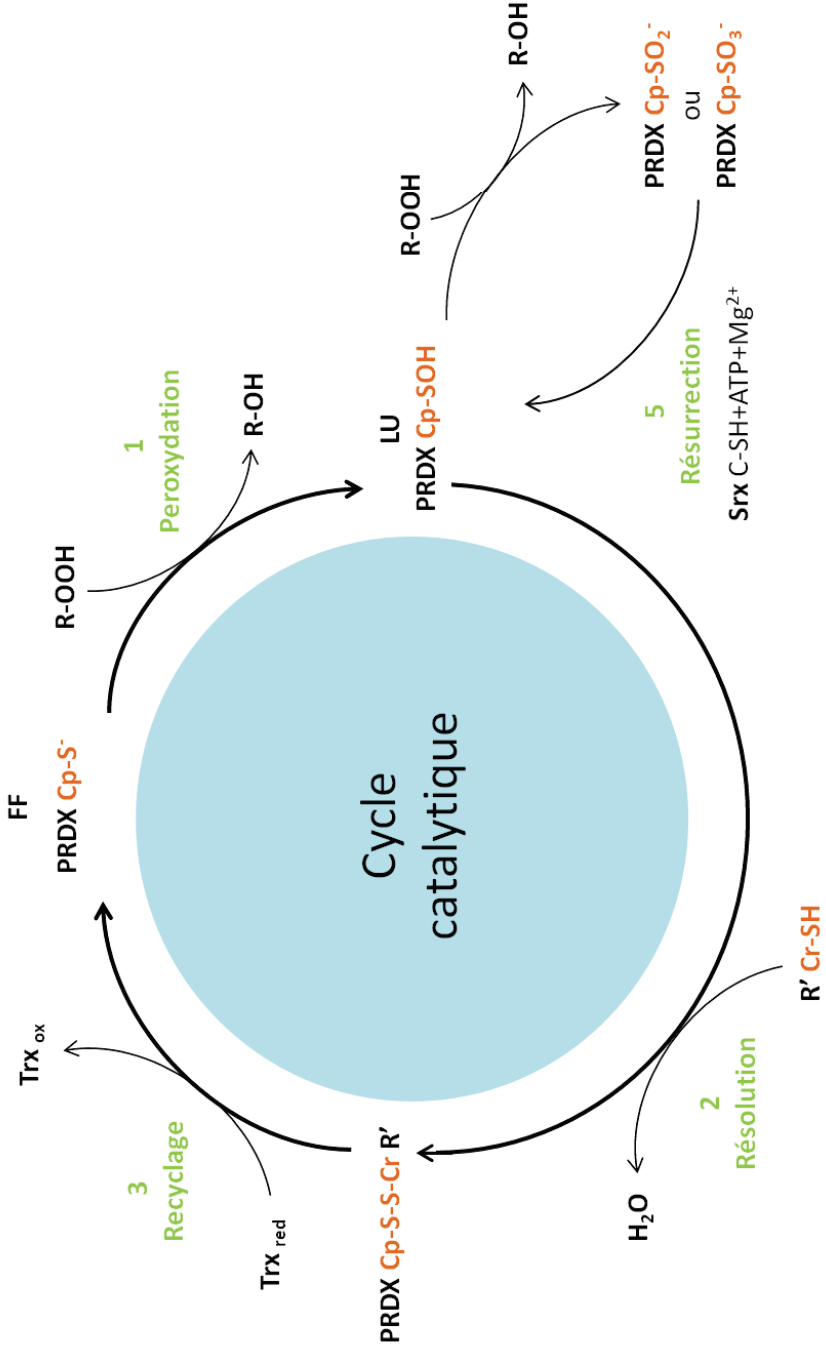


Figure 28. Schéma des 5 étapes du cycle catalytique de réduction des peroxydes par les PRDX.

L'étape 1 de peroxydation correspond à l'attaque nucléophile de Cp-S⁻ sur H₂O₂, menant à son oxydation en Cp-SOH. Lors de l'étape 2 de résolution il y a création d'un pont disulfure (intermoléculaire ou intramoléculaire selon la classe des PRDX) entre la Cp et la Cr. L'étape 3 de recyclage permet la réduction du pont disulfure pour reformer Cp-SOH afin de démarrer un nouveau cycle catalytique. L'étape possible de suroxydation s'effectue par l'attaque d'un deuxième (ou troisième peroxyde), menant à l'obtention d'un sulfinate Cp-SO₂⁻ (ou sulfonate Cp-SO₃⁻). L'enzyme ainsi inactivée nécessite l'action d'une enzyme : la sulfirédoxine (Srx), pour être régénérée. LU = *Locally Unfolded* et FF = *Fully Folded*.

Figure reproduite et modifiée à partir de Perkins et al. (2014), "Tuning of peroxidoredoxin catalysis for various physiological roles." *Biochemistry* **53** (49): 7693-7705.¹²⁰

Enfin, l'étape de recyclage va avoir pour but de réduire le pont disulfure et de reformer Cp-SOH afin de commencer un nouveau cycle, ce qui se fait *in vivo* par le système Trx/TrxR. Cette réduction peut être effectuée directement *in vitro* par différents réducteurs tels des mercaptans (thiols) comme le dithiothréitol (DTT) ou des phosphines comme le tris(2-carboxyéthyl)phosphine (TCEP). La protéine dans son état LU peut alors retrouver l'état FF afin de redémarrer un nouveau cycle catalytique.¹³⁰

Lorsque l'enzyme se trouve sous forme oxydée Cp-SOH l'étape de résolution n'est pas la seule réaction qui peut avoir lieu. En effet, une seconde molécule de peroxyde peut réagir avec Cp-SOH. Il y a alors une nouvelle étape d'oxydation de la cystéine qui se transforme en sulfinat Cp-SO₂⁻. Cette étape de suroxydation ne permet plus de participer au cycle catalytique. Pour les isoformes des PRDX à 2 cystéines, l'étape de « résurrection » est catalysée par la Srx et met en jeu le résidu cystéine conservé de la Srx, en présence d'ATP, d'ions Mg²⁺ et d'un donneur d'électron (thiol).

III. 1. 3 Modèle: la peroxyrédoxine 5 humaine

Parmi les 6 isoformes humaines, la hPRDX5 a été bien caractérisée par RMN au laboratoire. C'est un homodimère symétrique de 2x19 kDa, formé par une interface de type A (Figure 29). Dans chaque monomère qui compose le dimère, hPRDX5 possède 3 cystéines : une Cys conservée en position 47, et deux cystéines additionnelles localisées aux positions 72 et 151.^{91; 121} Plusieurs structures cristallographiques de hPRDX5 sont disponibles à partir de la PDB sous sa forme oxydée (PDB : 2VL2,2VL3)¹³¹ ou réduite (2VL9)¹³¹, ainsi que des structures dans lesquelles des ligands ont été caractérisés dans le site actif, notamment l'acide benzoïque (PDB : 1HD2, 1OC3, 2VL2)^{115; 131; 132} ou le DTT oxydé (PDB : 3MNG)¹⁰⁶ avec de bonnes résolutions (de l'ordre de 1.5Å).

Son expression et ses étapes de production et de purification ont d'ores et déjà été décrites et publiées par l'équipe de B. Knoop à l'Université Catholique de Louvain (Belgique), et bien maîtrisées au laboratoire de même que son marquage isotopique uniforme en ¹⁵N ou ¹³C.^{91; 115} L'environnement immédiat et accessible au solvant autour de Cp est majoritairement hydrophobe (Figure 29) avec les acides aminés Leu116, Phe120 et Phe240, Pro45 et 180 ou Ile119, formant une paroi d'interactions plutôt hydrophobes (vdW) comme discuté précédemment pour les hPRDX. On trouve aussi Arg127 mentionnée précédemment, orientée vers la cystéine afin de stabiliser la forme thiolate. Enfin ce site actif se situe à l'interface entre les deux chaînes de l'homodimère, et on notera la présence de Glu179, appartenant à l'autre chaîne de l'homodimère, et qui aura un rôle important dans la suite de l'étude.

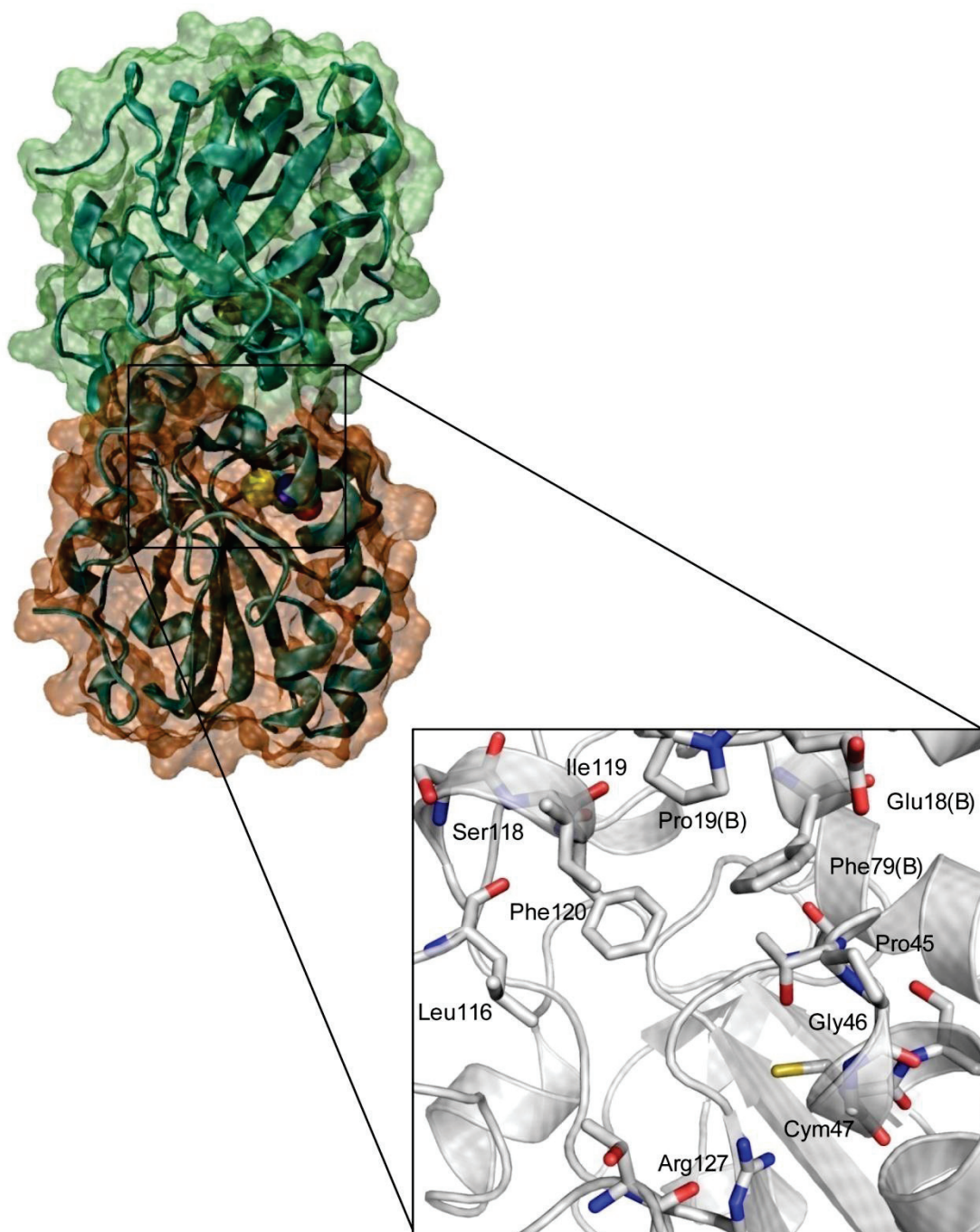


Figure 29. Visualisation de l'homodimère hPRDX5 et de son site actif.

Homodimère symétrique de hPRDX5 (3MNG.pdb) et sa surface colorée selon les chaînes A (orange) et B (vert). Structure du site actif centré autour de la cystéine catalytique Cys47 sous forme thiolate (Cp-S) Cym47. Visualisation agrandie des acides aminés constituant le site actif.

III. 2. Une implication biologique récemment proposée : l'ischémie

Les PRDX ont de nombreux rôles biologiques qui vont dépendre de leur localisation et de leur structure (comme vu en III.1.2). Leur quasi omniprésence dans les différents compartiments cellulaires et leur implication dans les phénomènes de stress oxydant vont en faire des cibles thérapeutiques potentielles dans le cas de diverses pathologies.¹⁰³ Par exemple, la modification de l'expression des hPRDXs a été reliée au développement de certains cancers même si leur rôle reste controversé.¹³³ Selon leur classe, les différentes isoformes de hPRDX ont été montrées comme induisant l'inhibition ou au contraire la promotion de cellules cancéreuses, dépendant de la localisation et du type de cancer. Si la surexpression des PRDX a tout d'abord été supposée comme cause de la résistance des cellules cancéreuses aux ROS via une augmentation de l'activité antioxydante, il est à présent proposé qu'elle puisse aussi agir comme un dispositif préventif de tumeurs cancéreuses via la détoxification de ces ROS.¹³³

Dans notre cas, nous allons voir qu'elles ont été plus particulièrement montrées il y a quelques années comme ayant un rôle médiateur dans l'inflammation post-ischémique cérébrale pouvant se produire à la suite d'un choc ischémique. En dehors de leur implication dans la promotion de cette inflammation, il a été par exemple observé que hPRDX2 subit une nitration d'un résidu tyrosine (formation de 3-nitrotyrosine) chez les patients se trouvant au stade précoce d'une maladie dégénérative : la maladie d'Alzheimer.¹²⁷

III. 2. 1 Les attaques vasculaires cérébrales

Un accident vasculaire cérébral (AVC) est lié à une perturbation de l'irrigation d'une zone du cerveau soit par obstruction d'un vaisseau sanguin (thrombose) soit par une rupture vasculaire hémorragique, conduisant à un déficit neurologique localisé. Les cellules nerveuses de cette zone sont endommagées et meurent, pouvant entraîner de nombreuses complications.¹³⁴ Dans le monde, on dénombre environ 15 millions de cas d'AVC par an sans aucune distinction de race ou d'ethnie.¹³⁵ Sur ces 15 millions de personnes un tiers décède, ce qui fait des AVC une source majeure de mortalité, mais aussi de handicap avec un autre tiers de personnes qui gardent des séquelles sur le long terme.¹³⁶ Plus précisément, en France les AVC touchent 130 000 personnes chaque année, ce qui en fait la 3^e cause de mortalité après les cancers et les cardiopathies, et la 1^{ère} cause de handicap physique.¹³⁷ C'est donc un sujet de santé publique primordial.

- Définition, facteurs de risques et complications

On distingue deux types d'AVC : ceux de nature ischémique et ceux de nature hémorragique (Figure 30).^{135; 138} Les AVC ischémiques ou infarctus cérébral sont les plus courants (> 85 % des AVC).

Ils sont dus à un arrêt du flux sanguin dans une artère provoqué par la présence d'un caillot de sang, qui peut provenir :

- d'un rétrécissement d'une artère (athéromatose) dû à une lésion interne, provoquant un caillot dans une artère principale, on parle de thrombose
- d'un caillot issu du cœur (embolie cardiaque)
- d'un caillot qui se serait formé directement dans une artère de petite taille à l'intérieur du cerveau

Les AVC hémorragiques sont eux liés à la rupture d'un vaisseau sanguin qui peut être favorisé par de nombreux facteurs tels que l'hypertension artérielle, le diabète, l'alcoolisme ou bien des troubles de la coagulation, mais ils peuvent aussi provenir d'une malformation d'un vaisseau ou une anomalie de coagulation sanguine. Ils sont beaucoup moins fréquents que les AVC ischémiques, à hauteur d'environ 15% des attaques.^{135; 139}

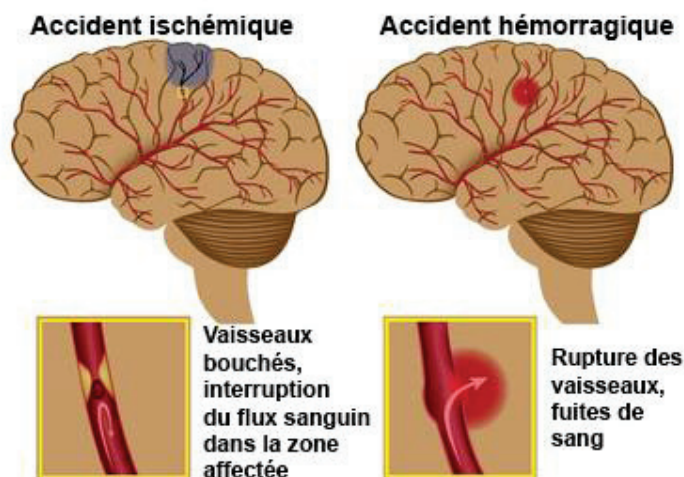


Figure 30. Représentation schématique des deux types d'accidents vasculaires cérébraux.

Différence entre l'accident vasculaire ischémique et hémorragique. Visualisation de la cause du choc au niveau des vaisseaux sanguins provenant des artères. L'accident ischémique est causé par un arrêt du flux sanguin conduisant à un manque d'irrigation de la zone du cerveau concernée, tandis que l'accident hémorragique est causé par la rupture d'un vaisseau menant à des fuites de sang. Image reproduite avec l'accord de Formation RCR PRO-VIE.

Une étude statistique mondiale, INTERSTROKE, effectuée en 2010 par des chercheurs de l'Université de Toronto a permis de mettre en évidence les principaux facteurs de risque rencontrés dans 80% des cas d'AVC. On retrouve les 5 principaux facteurs de risque suivants : l'hypertension artérielle, le tabagisme, l'obésité abdominale, une mauvaise alimentation et un manque d'activité physique.¹⁴⁰

Les AVC provoquent des dommages cérébraux qui peuvent se révéler parfois irréversibles. En effet, les cellules nerveuses ne se renouvèlent plus et finissent par mourir à cause de la privation d'oxygène, ce qui va entraîner des pertes fonctionnelles plus ou moins importantes selon la région touchée. Les complications neurologiques et médicales observées suite à un AVC peuvent être

diverses et dans la totalité des décès survenus par AVC, 20 à 50 % des décès sont dus à des complications post-AVC. Les œdèmes cérébraux et les transformations hémorragiques sont les deux principales sources de complications neurologiques qui surviennent rapidement après le choc ischémique et qui peuvent entraîner des conséquences lourdes à la fois sur le court et le long terme, tout comme les crises d'épilepsie ou bien la récurrence de chocs ischémiques.¹⁴¹ Enfin, les complications dues aux traitements administrés aux patients à la suite d'un AVC ne sont pas à négliger, tout comme l'intervention chirurgicale.¹⁴²

- Traitements

A la suite d'un AVC bénin, la prise de certains médicaments permet de pouvoir diminuer la possibilité de récurrence en agissant directement sur le facteur de risque responsable. Ainsi les antiagrégants plaquettaires tels que l'aspirine, le clopidogrel (Plavix®) ou le flurbiprofène (Froben®) et les anticoagulants comme la warfarine ou les héparines de faible poids moléculaire vont diminuer la possibilité de formation de caillots et donc diminuer le risque de thrombose au niveau du cœur et du cerveau.

Dans les cas d'hypertension artérielle, les antihypertenseurs antagonistes de l'angiotensine I (eprosartan (Teveten®)), losartan (Cozaar®) ont été étudiés de manière clinique (essai MOSES) dans le cadre de la prévention des AVC et montré comme réduisant de 25 % le risque de récurrence.¹³⁶ De plus, de nombreuses enzymes nommées activateurs de plasminogène (PAs) ont été étudiées comme potentiels agents thérapeutiques dans le traitement des chocs ischémiques, comme l'activateur tissulaire du plasminogène recombinant (rt-PA) ainsi que l'urokinase (u-PA), la streptokinase (SK), la reteplase (r-PA), la tenecteplase (TNK) et la staphylokinase (STK).¹⁴³

Il existe à ce jour un traitement de base des chocs ischémiques tiré de ces études qui consiste à administrer par intraveineuse le rt-PA qui va avoir pour action de dissoudre le caillot sanguin qui obstrue l'artère. Seulement les traitements utilisant un PAs ne sont efficaces que s'il est administré dans les 4h30 suivant l'AVC. De plus, si ce traitement permet de réduire les déficits neurologiques, il comporte néanmoins un risque d'hémorragie cérébrale.^{134; 144} Enfin, ce traitement appelé thrombolyse n'est actif que sur environ 35 % des patients, et se révèle complètement inefficace pour ceux présentant un thrombus riche en plaquettes liées entre elles via des récepteurs Gplb alpha (glycoprotéines) qui rend les thrombus résistants au tr-PA. Cette observation peut mener à de nouvelles stratégies thérapeutiques visant à rétablir la perméabilité artérielle.¹⁴⁵

Au lieu de chercher à restaurer la circulation sanguine pour oxygéner rapidement le cerveau, une autre solution vise plutôt à protéger les tissus encore viables jusqu'à ce qu'il puisse être de nouveau irrigué et donc ré-oxygéné. Pour cela, des chercheurs ont montré que la délivrance d'oxygène par un

masque facial permettait de prévenir quasi-complètement la perte neuronale ainsi que l'inflammation tissulaire dans le cas de modèles murins. Ce traitement pourrait être facilement transposable à l'homme et mis en œuvre dans la première heure suivant l'AVC.¹⁴⁶

Il en résulte que malgré de nombreux tests effectués au cours des dernières années, aucun traitement n'a réellement prouvé son efficacité et la recherche d'un traitement des chocs ischémiques reste une préoccupation majeure dans le milieu de la recherche médicale. Pour cela il est important de mieux comprendre les phénomènes qui ont lieu au cours de la cascade de réactions biologiques provoquées lors d'un choc ischémique, et au cours des heures suivant son apparition.¹⁴⁷

III. 2. 2 Le rôle des peroxyrédoxines dans l'inflammation post-ischémique

- L'inflammation post-ischémique cérébrale

L'accident vasculaire cérébral de nature ischémique s'accompagne d'une cascade d'évènements biologiques qui évoluent au cours du temps, et qui ont pour conséquence la destruction des parois cellulaires ainsi que la mort des cellules nerveuses localisées dans la zone de l'AVC. Ces processus destructeurs débutent immédiatement au cours de l'ischémie, tels que l'excitotoxicité, ou encore le stress oxydant, et peuvent se prolonger parfois plusieurs jours comme le phénomène de mort cellulaire : la nécrose, ou un phénomène inflammatoire qu'on appelle l'inflammation post-ischémique cérébrale.^{134;135;138}

L'une des conséquences immédiate de l'ischémie est l'augmentation de la concentration intracellulaire en calcium, et la production d'espèces réactives oxygénées (ROS). Ces espèces sont générées par le stress oxydant créé par les réactions radicalaires en chaîne de l'oxyde nitrique (NO) dont sa synthèse par l'oxyde nitrique synthase (NOS) augmente fortement entre 12 et 48h après l'ischémie.¹⁴⁸ L'augmentation de la production en NO associée à la forte concentration de calcium et la production d'un stress oxydant vont conduire à l'activation de facteurs de transcription (tel que NFκB) qui vont avoir pour conséquence l'expression de nombreuses cytokines pro-inflammatoires (inter leucine IL, facteur de nécrose tumorale TNF) et autres médiateurs de l'inflammation tels les chimiokines et les metalloprotéases.^{134;144;149} Ces médiateurs vont avoir à la fois un rôle neuroprotecteur en participant à la restructuration des tissus endommagés, mais aussi un rôle neurotoxique en aggravant les lésions ischémiques.¹⁴⁸

- Implication des peroxyrédoxines

De nombreuses protéines sont relâchées hors du compartiment cellulaire lors de la nécrose des cellules, qu'on appelle communément les DAMPs (motifs moléculaires associés aux dégâts

cellulaires) et qui vont aller interagir avec des récepteurs de type Toll (TLR) des macrophages. Ces DAMPs comprennent entre autres les protéines à haute mobilité de groupe 1 (HMGB1), les protéines de choc thermique (HSPs) et l'amyloïde β ($A\beta$).^{144,147}

Plus récemment, les PRDX ont été découvertes par Shichita *et al.* comme ayant aussi un rôle de DAMPs. En effet, les PRDX dans les cellules du cerveau sont surexprimées durant le stress ischémique, et libérées dans le compartiment extracellulaire lors de la nécrose de la cellule endommagée. Elles vont alors perdre leur caractère neuroprotecteur et aller se lier aux récepteurs TLR2 et TLR4 des macrophages qui vont initier la production de cytokines pro-inflammatoires (TNF α , IL-1 et IL-6) et autres médiateurs de l'inflammation, et ainsi déclencher la réponse inflammatoire (Figure 32).^{110; 138; 144; 147} Dans ce travail, des tests biologiques ont montré que hPRDX5 aurait la plus grande activité dans la production de cytokines inflammatoires. L'utilisation d'anticorps spécifiques à cette isoforme a donné lieu à la décroissance de l'expression de ces cytokines et l'atténuation des séquelles cérébrales pour des injections ayant lieu jusque 12h après le choc cérébral. De plus, la création d'anticorps spécifiques à différents motifs de l'isoforme 5 des peroxyrédoxines humaines a permis de mettre en évidence le motif 68 - 90 comme induisant la plus forte activité, et donc comme étant un cible thérapeutique potentielle (Figure 31). Ce motif comprend une hélice α exposée au solvant ainsi qu'un brin β enfoui dans la structure. Le site récepteur formé autour de la Cp étant entouré par ce motif, un inhibiteur affiné pour le site récepteur pourrait empêcher l'accès à ce motif 68-90 et donc affecter la production de cytokines inflammatoires.

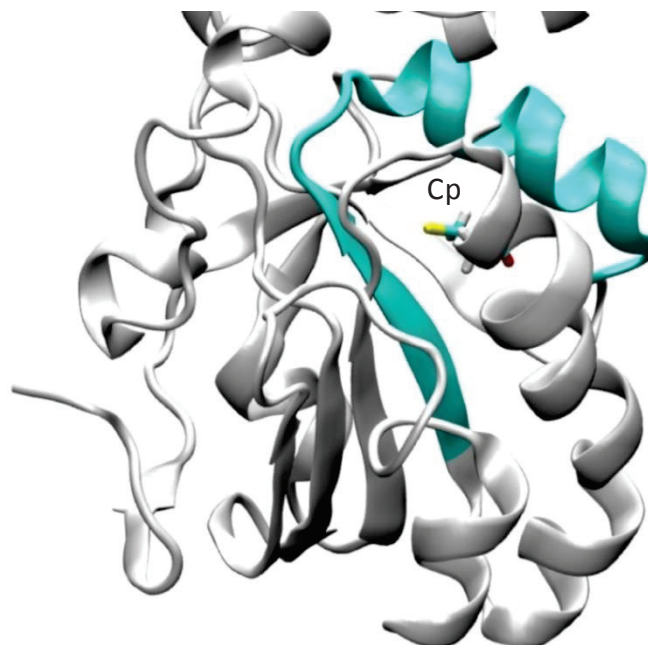


Figure 31. Visualisation du motif 68-90 de hPRDX5 impliqué dans la production de cytokines inflammatoires. Structure de hPRDX5 sous forme réduite avec le motif 68-90 montré comme ayant la plus grande implication, coloré en cyan, et le résidu Cp (Cys47) coloré selon la nature des atomes.

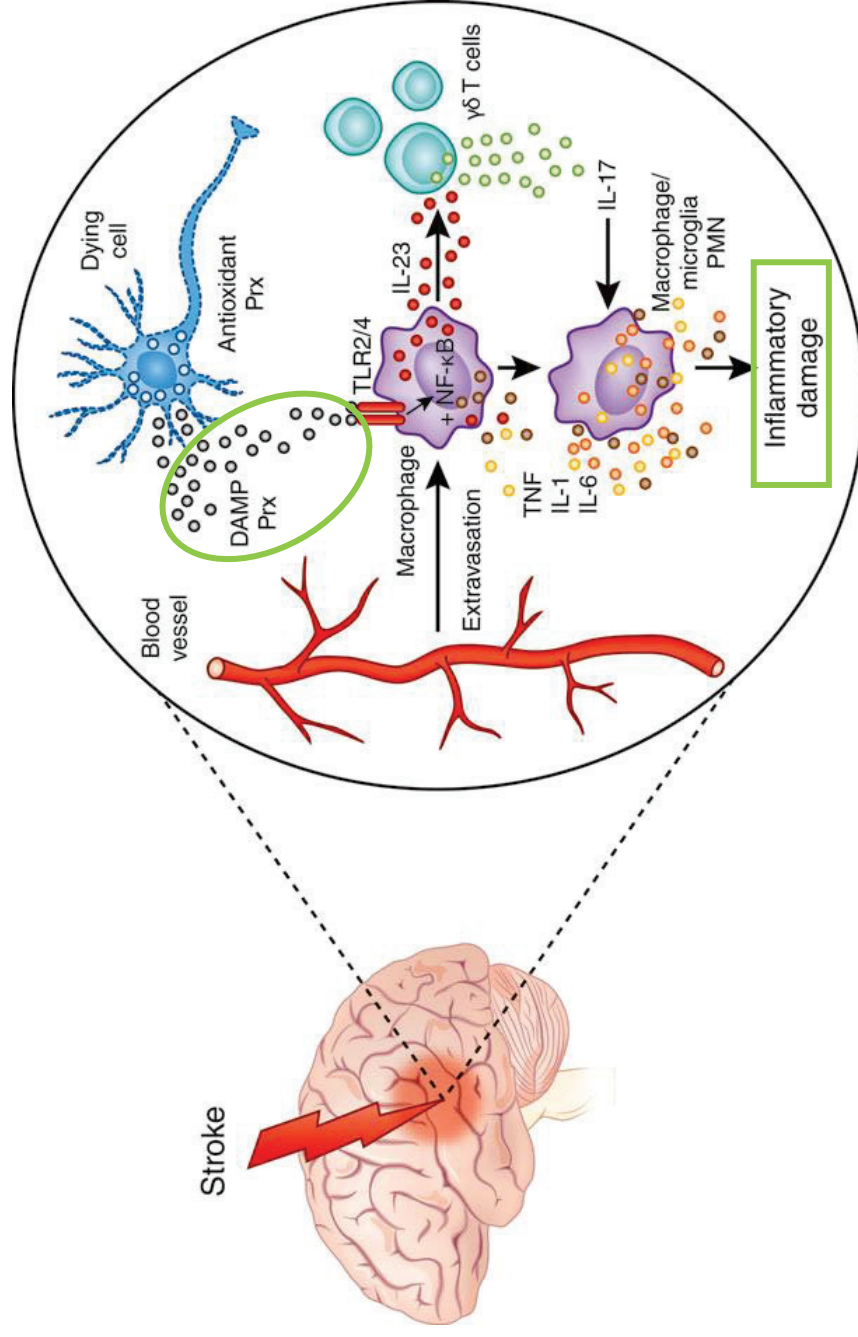


Figure 32. Visualisation de la cascade ischémique survenant après un AVC.

Lors d'un choc ischémique, les neurones nécrosés vont libérer leur matériel cellulaire, dont les peroxyradicaux qui ont été identifiés comme pouvant se lier aux récepteurs des macrophages et déclencher une inflammation post-ischémique. Image reproduite de Garcia-Bonilla et al. 2012 avec l'accord de Nature Publishing.

III. 3. Inhibition des peroxyrédoxines

Les inhibiteurs sont des composés capables de ralentir la réaction catalysée par l'enzyme lorsqu'ils sont mis en présence de celle-ci. L'inhibition des enzymes joue un rôle capital dans le contrôle des mécanismes biologiques puisqu'un inhibiteur peut affecter la fixation de substrats sur le site actif et/ou modifier les stabilisations des états de transition. Nombre de médicaments ont des propriétés inhibitrices pour des enzymes, c'est pourquoi leur découverte et leur amélioration est un domaine de recherche actif en biochimie et pharmacologie.

En première approche, un inhibiteur est défini à travers la détermination de sa concentration inhibitrice médiane (IC_{50}) soit la concentration de cet inhibiteur nécessaire pour inhiber la moitié de l'activité de l'enzyme, que l'on peut classiquement déduire d'expériences de cinétique enzymatique.¹⁵⁰ La protéine d'intérêt de cette étude, hPRDX5, possède 2 sites d'interaction, un sur chaque chaîne de l'homodimère. Les enzymes possédant plusieurs sites de fixation peuvent présenter un comportement coopératif (allostérie), c'est-à-dire que la liaison d'un ligand dans l'un des sites peut entraîner la modification de l'affinité pour le second site, via un changement conformationnel. Le comportement michaelien de hPRDX5 ayant été préalablement démontré par une autre doctorante de l'équipe, nous allons plus particulièrement voir ci-dessous les enzymes présentant un modèle hyperbolique de Michaelis-Menten.

III. 3. 1 Cinétique enzymatique

Nous définirons tout d'abord la cinétique dite michaelienne et l'équation de Michaelis-Menten pour ces enzymes, nous présenterons ensuite les différentes expériences de cinétique enzymatique pour la mesure de l'inhibition, et présenterons les différents mécanismes d'inhibition réversible existants.^{151; 152; 153}

- Michaelis-Menten

Comme présenté en première partie d'introduction (I.1.1.4), la formation d'un complexe PL ou dans ce cas particulier d'une enzyme (E) avec un substrat (S) donnant un complexe enzyme-substrat (ES), passe par un état de complexation réversible (état de transition). Plus précisément, toutes les réactions enzymatiques michaeliennes démarrent par la formation d'un complexe binaire mettant en jeu l'enzyme et son substrat complexé de manière non covalente, avant transformation en produit. Ce complexe est appelé communément le complexe de Michaelis. Sa formation se fait de manière réversible ce qui permet d'écrire l'équation de réaction suivante :



Avec P le produit de la réaction, k_1 et k_{-1} les constantes de vitesse de l'équilibre réversible, et k_{cat} la constante catalytique de la réaction qui caractérise l'efficacité catalytique de l'enzyme. En considérant un état stationnaire correspondant à l'équilibre entre la formation et la disparition du complexe ES dans lequel la quantité de produit P est négligeable, et la vitesse de la réaction reste constante, l'équation de Michaelis-Menten **30** donne la forme générale de l'équation de vitesse d'une réaction enzymatique (dans le cas d'un état stationnaire, cette vitesse est la vitesse initiale v_i) :

$$v_i = \frac{k_{cat}[E]_0[S]}{K_M + [S]} = \frac{v_{max}[S]}{K_M + [S]} \quad [31]$$

$$K_M = \frac{k_{-1} + k_{cat}}{k_1} \quad [32]$$

Avec K_M la constante de Michaelis, k_{cat} la constante de vitesse du premier ordre définissant la capacité du complexe enzyme-substrat à former le produit de la réaction, et $V_{max} = k_{cat}[E]_0$ la vitesse limite, ou vitesse maximum, dépendante de la concentration initiale en enzyme.¹⁵²

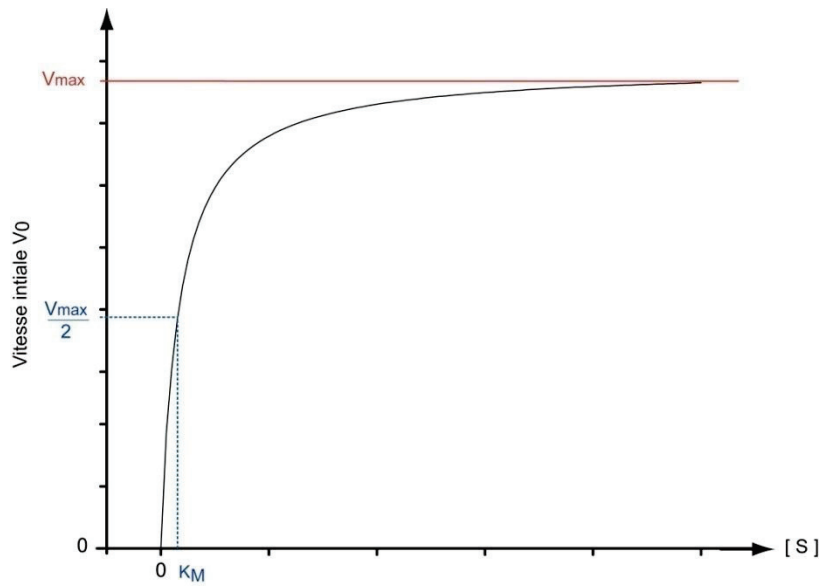
K_M est donc une mesure relative d'affinité de l'enzyme pour son substrat. Elle représente la concentration en substrat pour laquelle la vitesse est égale à la moitié de la vitesse limite. D'après l'équation **31** de Michaelis, $v_i = \frac{1}{2} V_{max}$ lorsque $[S] = K_M$. Il est alors possible de déterminer graphiquement V_{max} , et d'en déduire la valeur de K_M par tracé d'un graphique représentant les vitesses de réaction en fonction de la concentration en substrat $v_i = f([S])$ (Figure 32.A). Une autre représentation consiste à tracer ce graphique de manière linéaire en traçant $1/v_i = f(1/[S])$, les intersections de la droite avec les axes permettant de remonter aux valeurs de K_M et V_{max} (Figure 33.B).¹⁵⁰

L'équilibre **30** est quantifiable par la valeur de la constante de dissociation du complexe ES , appelée K_D (ou dans le cas spécifique du complexe ES , K_S) exprimée en unité molaire :

$$K_S = \frac{k_{-1}}{k_1} \quad [33]$$

De plus la réversibilité de cet équilibre met en jeu les constantes de vitesse d'association et de dissociation du complexe : k_1 (k_{on}) et k_{-1} (k_{off}). Ainsi, si la réaction chimique est lente en comparaison de la dissociation du complexe ES , alors k_{cat} est négligeable vis-à-vis de k_{-1} , ce qui implique dans ce cas que $K_M = K_S$. C'est ce que nous considérerons dans le reste du manuscrit puisque dans le cas de notre étude, il a été préalablement vérifié que $k_{cat} \ll k_{-1}$.¹⁵³

A)



B)

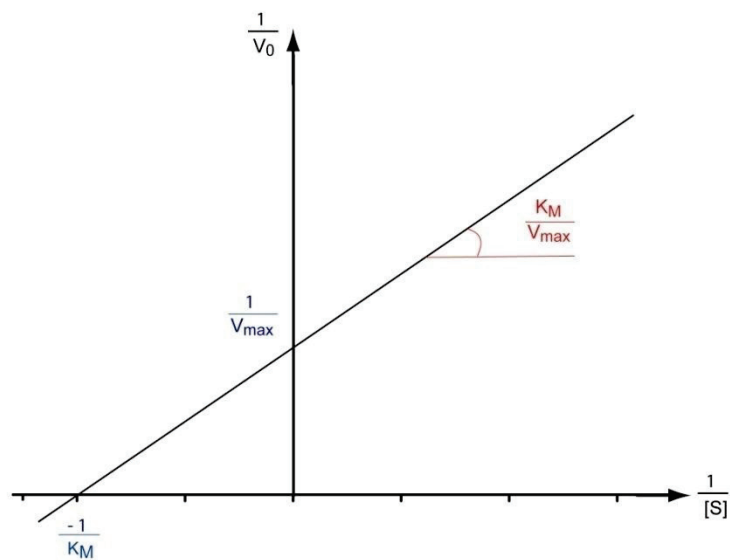


Figure 33. Représentations graphiques des v_i en fonction de la concentration en substrat pour des enzymes michaeliennes, et détermination graphique de K_M et V_{max} .

A) Représentation classique hyperbolique: $v_0 = f([S])$, avec $[S] = K_M$ à $V_{max}/2$. B) Représentation linéaire de Lineweaver-Burk (1934) : $1/v_0 = f(1/[S])$ où l'ordonnée à l'origine correspond à $1/V_{max}$, et K_M est déterminé par $-1/K_M$, la valeur où la droite coupe l'axe des abscisses.

III. 3. 2 Mesure de l'activité enzymatique

Afin de mesurer l'activité inhibitrice d'un ligand sur une enzyme, des expériences de cinétique enzymatique *in vitro* peuvent être effectuées. Ces expériences sont mises en place en utilisant des techniques telles que des mesures d'absorbance ou de fluorescence. L'activité enzymatique est déterminée par calcul de la vitesse initiale de réaction, en présence et en absence d'inhibiteur. Pour cela il faut tout d'abord déterminer les conditions optimales de la réaction catalytique (concentration des différentes enzymes du système) et pour une observation aisée et précise des v_i . Il est alors possible d'identifier le potentiel inhibiteur de différents ligands à travers la mesure de la concentration inhibitrice moyenne (IC_{50}). Cette concentration est mesurée par des expériences de cinétique enzymatique dans lesquelles la protéine est mise en présence d'une quantité de substrat équivalente au K_M , et des concentrations croissantes de ligand. Ainsi en gardant les mêmes conditions d'expérience, il est possible de discriminer les différents ligands non pas uniquement sur leur affinité, mais sur leur capacité d'inhibition. Ces mesures de cinétique enzymatique permettent la mesure de l'activité inhibitrice de potentiels inhibiteurs, mais aussi de pouvoir déterminer le mécanisme d'inhibition de ces inhibiteurs envers l'enzyme, ce que nous allons maintenant considérer rapidement.^{150; 153}

III. 3. 3 Le mécanisme d'inhibition enzymatique

- Les différents mécanismes d'inhibition enzymatique

L'inhibition enzymatique peut se faire de manière réversible ou irréversible. Elle est dite irréversible lorsqu'il est impossible de retrouver l'activité initiale après exposition aux inhibiteurs. Soit les inhibiteurs se lient à l'enzyme par des liaisons chimiques, soit l'enzyme est dénaturée. Dans le cas d'une inhibition réversible, il est possible de lever l'inhibition en éliminant l'inhibiteur du milieu réactionnel par dialyse ou dilution, tandis que dans le cas d'une inhibition irréversible, il y a modification covalente d'une fonction de l'enzyme qui ne peut pas être levée. Cette réversibilité peut être étudiée par le biais de l'étude des constantes de vitesses.^{150; 151} Une fois la réversibilité de l'inhibition déterminée, on peut déterminer le mécanisme avec lequel l'inhibiteur se lie à l'enzyme. Il existe de nombreux modèles associés à l'inhibition réversible.

Les inhibitions auxquelles nous allons plus spécifiquement nous intéresser sont les inhibitions dites linéaires, qui regroupent les 3 principaux mécanismes d'inhibitions de base : l'inhibition compétitive, l'inhibition non-compétitive, et l'inhibition incompétitive. Ces mécanismes définis dans la Figure 34 ont des effets spécifiques sur les valeurs de K_M et de V_{max} .^{151; 152}

L'inhibition compétitive

L'inhibition compétitive est caractérisée par la compétition du substrat et de l'inhibiteur pour l'enzyme libre. Ainsi, si le substrat est lié à son enzyme, l'inhibiteur ne peut s'y lier, et vice-versa. Dans ce type de mécanisme, la liaison de l'inhibiteur au site de l'enzyme ne permet aucune conversion en produit (Figure 34). On peut noter que l'inhibiteur compétitif possède souvent une ressemblance structurale avec le substrat originel de l'enzyme. Dans ce type de mécanisme, le V_{max} n'est pas affecté par la concentration en inhibiteur $[I]$, contrairement au K_M qui augmente avec $[I]$.

L'inhibition incompétitive

Dans un mécanisme d'inhibition incompétitive, au contraire de l'inhibition compétitive, l'inhibiteur ne se fixe jamais à l'enzyme libre mais uniquement au complexe enzyme-substrat ES afin d'inhiber la formation des produits (Figure 34). Il y a le plus souvent modification de la conformation de l'enzyme grâce à la liaison du substrat, permettant ainsi à l'inhibiteur de venir se lier. On reconnaît ce type de mécanisme lorsque l'on observe une décroissance du V_{max} et du K_M pour $[I]$ qui augmente.

L'inhibition non-compétitive

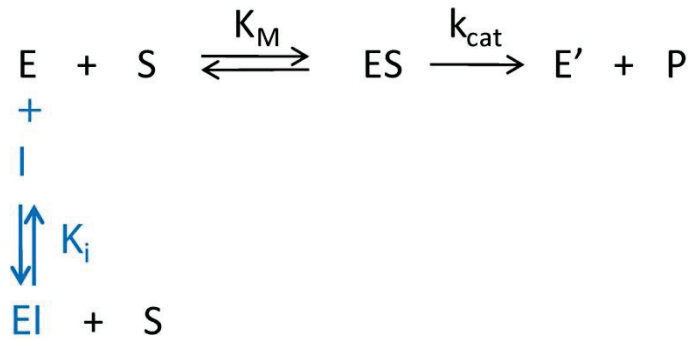
Dans un mécanisme d'inhibition dite non-compétitive, l'inhibiteur n'induit aucun effet sur la liaison du substrat. Il y a liaison du substrat et de l'inhibiteur de manière réversible et indépendante, sur des sites de l'enzyme pouvant être différents. Ainsi, l'inhibiteur peut se lier à l'enzyme libre pour former un complexe binaire EI , comme au complexe ES formant alors un complexe ternaire ESI avec des constantes d'inhibitions distinctes, respectivement K_i pour le complexe EI , et K'_i (αK_i) pour le complexe ESI (Figure 34). Il en résulte que selon les valeurs relatives de K_i et K'_i , plusieurs mécanismes peuvent être définis.

Le modèle classique d'inhibition non-compétitive est le modèle dans lequel $K_i = K'_i$, et pour lequel on observe une diminution du V_{max} lorsque $[I]$ augmente, sans impacter la valeur du K_M . Il existe ensuite un modèle mixte dans lequel on distingue 2 cas possibles. Soit l'inhibiteur tend à se lier préférentiellement à l'enzyme libre ($K_i > K'_i$) soit au complexe ES ($K_i < K'_i$). Lorsque $K_i > K'_i$, K_M et V_{max} vont diminuer avec l'augmentation de $[I]$, tandis que lorsque $K_i < K'_i$, V_{max} augmente avec $[I]$ mais K_M diminue. Enfin, un modèle partiel-mixte non compétitif, non linéaire, existe lorsqu'il y a une conversion possible à partir du complexe ESI avec une constante catalytique k_{cat}' (Figure 34).

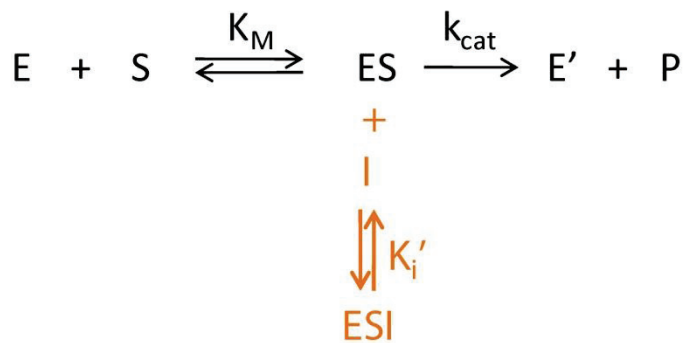
- La détermination du mécanisme d'inhibition

La détermination expérimentale du mécanisme d'inhibition peut se faire par des études cinétiques, tout comme pour la mesure de l' IC_{50} .

Mécanisme Compétitif



Mécanisme Incompétitif



Mécanismes Non-compétitif

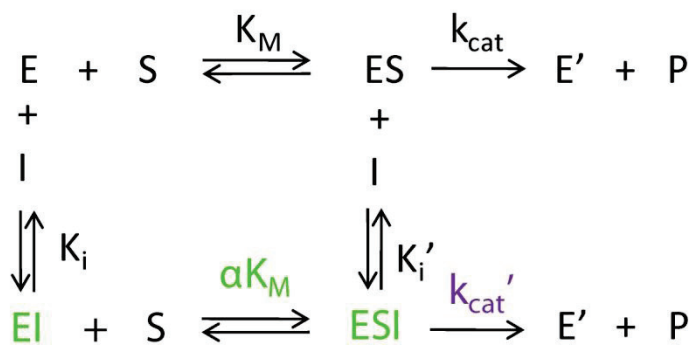


Figure 34. Schéma des différents mécanismes d'inhibition réversible présentés.

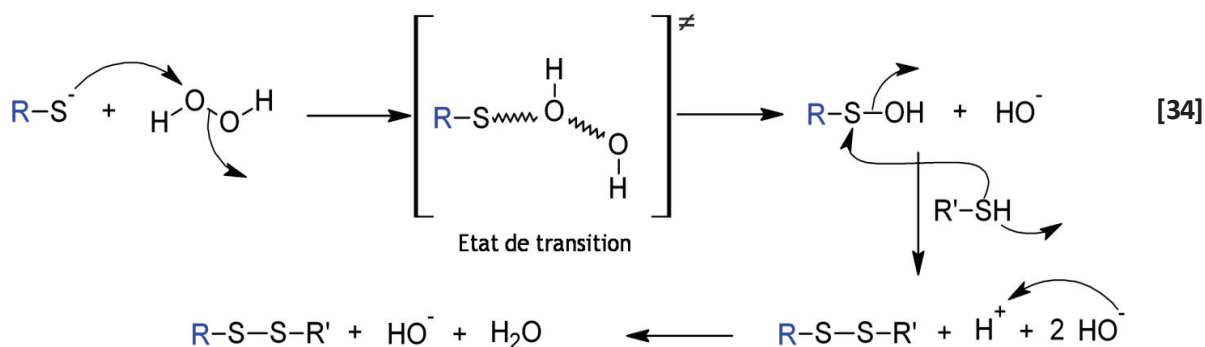
Le mécanisme compétitif dans lequel l'inhibiteur I se liera préférentiellement à l'enzyme libre E , au contraire du mécanisme incompetitif pour lequel I se lie préférentiellement au complexe enzyme-substrat ES . Les mécanismes non-compétitifs pour lesquels I se lie à la fois à E et au complexe ES . Différents mécanismes non-compétitifs existent selon la valeur des constantes d'inhibitions K_i et K_i' (αK_i), un mécanisme classique pour $K_i = K_i'$, deux mécanismes mixtes pour $K_i > K_i'$ et $K_i < K_i'$, et un mécanisme partiel-mixte pour k_{cat}' non nulle, c'est à dire lorsqu'il ya conversion de l'enzyme à partir du complexe ternaire ESI avec une constante différente de k_{cat} .

Seulement, afin de déterminer le mécanisme d'inhibition, il faut effectuer la mesure de la vitesse initiale de réaction à concentration constante en inhibiteur et par des concentrations croissantes de substrat. Il faut ensuite répéter l'opération pour différentes concentrations en inhibiteur, choisies inférieures, égale et supérieures à la valeur de l' IC_{50} de l'inhibiteur considéré. Le mécanisme d'inhibition est ensuite déterminé par tracé des vitesses initiales en fonction de la concentration en substrat, pour les différentes concentrations en inhibiteur $[I]$, en fonction de l'effet causé par l'augmentation de $[I]$ sur les valeurs de V_{max} et K_M .^{150; 151; 153}

III. 3. 4 La réaction enzymatique des hPRDXs

Nous avons vu que la famille des hPRDX est connue pour catalyser la réaction de réduction de divers peroxydes, tels que le peroxyde d'hydrogène dont la régulation à l'intérieur de la cellule est effectuée grâce à ces enzymes.^{103; 106}

La réaction **34** catalysée par hPRDX, conduit à la réduction des peroxydes via la Cp, et la formation d'un pont disulfure et de molécule d'eau. Le thiolate, nucléophile, va attaquer l'un des deux atomes d'oxygène de H_2O_2 .



Cette réaction va mener à la formation d'un acide sulfénique et d'un ion hydroxyde, en passant par un état de transition (ET) dans lequel la liaison O-O commence à se dissocier, et la liaison S-O commence à se former. L'attaque d'un second groupement thiol sur la PRDX oxydée amène la formation d'un pont disulfure, et la libération d'une molécule d'eau.^{103; 106}

L'état de transition de la réaction catalysée par les hPRDX a déjà été modélisé en 2014 par QM/MM sur un jeu réduit d'atomes (Figure 35).¹⁵⁴ Hall et al. en 2010 avaient interprété qualitativement l'ET comme un réseau d'interactions non covalentes entre Cp-S et H_2O_2 mettant en jeu les acides aminés présents dans l'entourage de la Cp, conservés, tels que Arg127, Thr44, Pro40 ou His51.¹⁰⁶

Les calculs de mécanique quantique publiés en 2014 par Portillo-Ledesma et al. viennent confirmer le rôle central de certains de ces résidus (Thr44, Arg127) dans la catalyse de cette réaction.

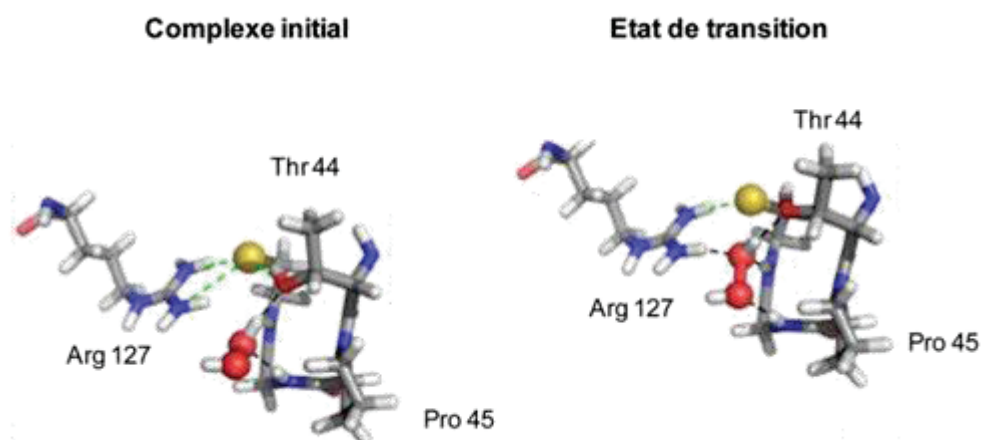


Figure 35. Structure du complexe réactionnel initial, et de l'état de transition dans la réaction enzymatique de réduction de H_2O_2 catalysée par hPRDX.

Structures calculées par mécanique quantique/mécanique moléculaire (QM/MM) du complexe initial et de l'état de transition de la réaction de réduction de H_2O_2 catalysée par hPRDX5. Entre le complexe initial et l'état de transition, on observe un allongement de la liaison O-O, parallèlement à une diminution de la liaison S-O. La région de la protéine utilisée pour les calculs de QM est représentée en bâtons. Le reste de la protéine ayant été modélisés en MM. Le réseau de liaisons H formées entre Cp-S, H_2O_2 et les résidus de hPRDX5 est représenté en lignes pointillées. Figure reproduite depuis Portillo-Ledesma et al. 2014 avec l'autorisation de American Chemical Society.

La connaissance de l'ET est importante car il est bien connu que les analogues de l'ET sont en général de très bons inhibiteurs.¹⁵⁵ De ces structures, il ressort que toute molécule ayant deux groupements OH proches comme dans l'ET, pourrait posséder une analogie structurale avec l'ET. De manière générale les molécules de type diol-1,2 pourraient donc être considérées comme ayant une ressemblance structurale avec l'ET du fait de la présence des deux groupements -OH adjacents. Cette hypothèse pourrait corrélérer le fait que le DTT a été trouvé par Hall et al. comme ligand dans le site actif de hPRDX5 lors des expériences de cristallisation.¹⁰⁶

III. 3. 5 Inhibiteurs connus de peroxyrédoxines humaines

Dans la littérature, plusieurs inhibiteurs irréversibles de hPRDX ont déjà été publiés. La plupart d'entre eux ont été démontrés comme inhibant hPRDX1 comme AMRI-59, et H7¹⁵⁶; ¹⁵⁷ puisque hPRDX1 est particulièrement étudiée du fait de sa surexpression dans le cas de cancers.¹⁵¹ ¹⁵² D'autres composés issus de produits naturels comme des terpènes : le célastrol et le triptolide, ou encore une lactone : la withaférine A, ont aussi été montrés comme étant des inhibiteurs de cette

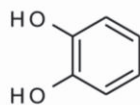
isoforme.^{158; 159} Comme énoncé précédemment, les hPRDX possèdent des structures quaternaires différentes, et des rôles cellulaires différents. Il en résulte aussi une possible sélectivité dû aux différences structurales. Ainsi H7 a été montré comme se liant directement au site actif de hPRDX1, mais aucune interaction n'a été observée pour les isoformes 2 à 5.¹⁵⁷ Au contraire, d'autres inhibiteurs ont été démontré comme ayant action sur plusieurs isoformes de hPRDX comme la conoïdine A^{160; 161} qui a été prouvée comme inhibant hPRDX1 mais aussi hPRDX2 et hPRDX4, de même que l'adénanthine, qui possède elle aussi une action inhibitrice à la fois sur hPRDX1 et hPRDX2.¹⁶²

- Les catéchols comme inhibiteurs de hPRDX5

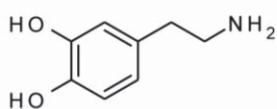
Nous avons vu en début de cette partie que le criblage RMN effectué dans notre laboratoire en 2010, avait permis d'identifier par STD l'affinité de 9 fragments *hits* pour hPRDX5. Puis, des expériences HSQC ont permis de prouver que ces fragments induisaient effectivement des perturbations de déplacement chimique (CSP) significatives, principalement localisées au niveau du site récepteur de hPRDX5. Cette étude a permis de mettre en évidence le *o*-pyrocatechol (plus simplement appelé catéchol) et son dérivé 4-méthylcatéchol parmi ces fragments, inhibant hPRDX5 avec des IC_{50} respectifs de $105 \pm 9 \mu\text{M}$ et $26 \pm 2 \mu\text{M}$.⁹¹

Le catéchol (CAT) ou 1,2-dihydroxybenzene de formule brute $\text{C}_6\text{H}_6\text{O}_2$ est un noyau aromatique à 6 carbones portant deux fonctions hydroxyles $-\text{OH}$ en ortho l'une de l'autre (Figure 36). Le catéchol est un des isomères du benzènediol (dihydroxybenzène), présent dans le squelette de nombreuses molécules naturelles. Tout d'abord dans certains neurotransmetteurs comme l'adrénaline, la noradrénaline ou bien la dopamine par exemple, appelés catécholamines, qui jouent aussi le rôle d'hormone dans la circulation sanguine (Figure 36).¹⁶³ On trouve aussi le catéchol dans la structure de certains polyphénols naturels comme la quercétine, ou la catéchine dont il fut isolé pour la première fois en 1839 (Figure 36). Leur toxicité est cependant controversée. Les catéchols étant présents à la fois dans notre corps, et dans de nombreux aliments, leur possible toxicité ne semble pas avoir d'impact sur la santé. De plus, des études ont montrées que les catéchols pouvaient avoir des effets anti-inflammatoires et neuro-protecteurs en jouant un rôle dans la suppression des cytokines inflammatoires.¹⁶⁴ Nous avons centré notre étude sur l'étude de ce motif catéchol, et de ces dérivés. Nous nous sommes fixés comme but de comprendre les détails de l'affinité des catéchols pour les hPRDX. Puis, sur ces bases, de vérifier et d'en améliorer l'inhibition, et d'en déterminer le mécanisme d'inhibition.

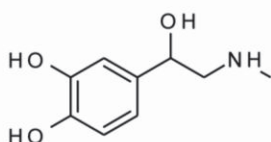
Catéchol



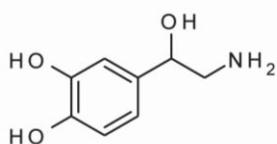
Neurotransmetteurs



Dopamine

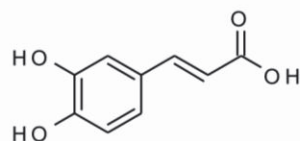


Adrénaline

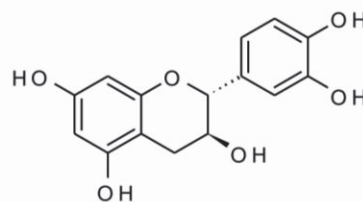


Noradrénaline

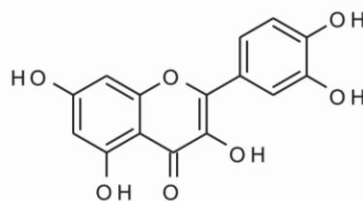
Polyphénols



Acide caféique



Catéchine



Quercétine

Figure 36. Structures du noyau *o*-pyrocatechol et de molécules naturelles possédant son motif.

Structure du catechol, et structures de neurotransmetteurs catécholamines : adrénaline, dopamine, noradrénaline, et de polyphénols : quercétine, catéchine, acide caféique.

CHAPITRE II

MATERIEL ET METHODES

I. PROTOCOLES DE MODELISATION

Toutes les simulations ont été effectuées grâce au logiciel ACEMD de la société Acellera⁶⁰, en utilisant le champ de force AMBER99SB-ILDN,⁴⁰ et ont été visualisées à l'aide du programme VMD (*Visual Molecular Dynamics*).¹⁶⁵ Les représentations graphiques des résultats de l'analyse des simulations ont été effectuées avec Gnuplot 4.4.¹⁶⁶ Toutes les simulations de dynamique moléculaire ont été réalisées à 300K sur un système (Metrocubo[®] Acellera) équipé de 4 GPU GTX 780 TI (NVIDIA[®]).

I. 1. Modélisation d'un système protéine-ligand

Tous les systèmes protéine-ligand ont été préparés à partir de la structure cristallographique 3MNG.pdb de hPRDX5 disponible sur la PDB,¹⁰⁶ et dont le dimère a été extrait des symétries cristallographiques à l'aide du logiciel PyMol (The PyMOL Molecular Graphics System, Version 1.4.1 Schrödinger, LLC) et la cystéine catalytique modélisée sous forme de thiolate. Les ligands ont été construits avec le logiciel Avogadro¹⁶⁷ et leurs structures minimisées en énergie. La suite d'outils Antechamber¹⁶⁸ a ensuite été utilisée afin de générer les fichiers associés au ligand, nécessaires à la réalisation de la dynamique dans le champ de force généralisé d'AMBER (GAFF)¹⁶⁹ : un fichier de charge (.prepi) et un fichier comprenant les éventuels paramètres de champ de force manquants associés à GAFF (utilitaire « parmchk »). Les fichiers .pdb du(des) ligand(s) et de la protéine ont été mis en commun afin de créer un fichier unique structure.pdb du système. Toujours en utilisant Antechamber, le système a ensuite été solvate et ionisé grâce au programme « Leap ». Le programme solvate le système par l'addition de molécules d'eau formant une boîte autour de cette structure dont la taille varie selon la taille du système modélisé et le ionise ensuite par l'ajout d'ions Na⁺ et Cl⁻ en fonction de la salinité, du pH et de la force ionique choisie en amont. Les molécules de solvant ont été modélisées de manière explicite selon le modèle TIP3P (voir I.3.2 du chapitre d'introduction) pour tous les systèmes.⁵⁵

Les acides aminés ont été modélisés à pH = 7.4 pour mimer au mieux les conditions expérimentales et physiologiques. Leur état d'ionisation a été défini à ce pH sans tenir compte de leur

environnement individuel excepté pour la cystéine catalytique (Cys47) qui a été modélisée sous forme thiolate S- (Cys47) due à la présence de Arg127 qui abaisse son pKa.

Avant tout lancement de dynamique moléculaire, les systèmes ont subi une phase d'équilibration-relaxation durant laquelle tous les atomes lourds du système (protéine et ligand) sont contraints localement à l'aide d'un potentiel harmonique afin d'éviter la réorientation du système. Ce processus se fait en 3 étapes :

- une courte période NVT (500 étapes de 4 fs) durant laquelle les contraintes locales appliquées sur les atomes lourds du squelette sont brutalement diminuées de 10 à 1 kcal.mol⁻¹.Å⁻².
- une première période NPT (nommée NPT1) de 1 ns à pression constante de 1 atm (101325 Pa), durant laquelle les contraintes sont graduellement diminuées de 1 à 0.1 kcal.mol⁻¹.Å⁻².
- une seconde période NPT (nommée NPT2) de 4 ns à pression constante de 1atm durant laquelle les contraintes initiales sont réappliquées, le volume de la boîte d'eau est laissé à relaxer.

Le système ainsi équilibré et relaxé en énergie est ensuite enregistré sous un nouveau fichier structure.pdb, qui sera utilisé dans le lancement de dynamiques moléculaires.

I. 2. Simulation de dynamique non biaisée

Les systèmes simulés en dynamique non biaisée NVT ont été construits comme expliqué au paragraphe précédent, en présence d'un excès de ligand. Pour ces systèmes, 6 molécules de ligand ont été modélisées autour de la protéine, à une distance < 30 Å des sites actifs de l'homodimère. Une molécule ligand connue peut aussi être modélisée à l'intérieur du site actif, telle que la molécule de DTT oxydée modélisée comme nous le verrons en partie I.1.2 du chapitre de résultats (p.105). Une fois le système construit et relaxé, la dynamique sans contrainte est lancée sur une durée initiale de 500 ns. Cette durée a été définie afin de visualiser au minimum un évènement de liaison pour des affinités comprises en le micromolaire et le millimolaire. Cette haute concentration en ligands permet d'augmenter la probabilité d'observer un évènement d'association.

I. 3. Simulation de FM

Pour les simulations de FM, le système obtenu après relaxation a été contraint localement par 3 atomes lourds de la protéine choisis en amont à la base de la protéine (20 kcal.mol⁻¹.Å⁻²). Le potentiel externe en forme d'entonnoir a ensuite été défini à l'aide d'un script qui nous a été fourni par notre collaborateur V. Limongelli, et visualisé dans VMD. Ce potentiel dépend des coordonnées des atomes de la protéine, et est à redéfinir avant chaque simulation.

Les coordonnées du potentiel externe, et les CV choisies pour le lancement de la FM ont ensuite été définies comme décrit en partie I.1.3 du chapitre de résultats (p.105-107) dans le logiciel Plumed 1.0 qui a été utilisé pour toutes les simulations de FM.⁶⁴ La simulation de FM a ensuite été réalisée sur une durée initiale de 500 ns afin de permettre une comparaison directe des résultats obtenus pour différents ligands. Cette durée peut être modulée et prolongée si nécessaire afin d'atteindre la convergence de l'énergie déposée.

I. 4. Modélisation d'un complexe PL en présence de substrat (H₂O₂)

Pour simuler un complexe Protéine-Substrat-Ligand par dynamique moléculaire, les coordonnées du substrat sont ajoutées au fichier de structure du système PL, qui est ensuite modélisé et équilibré comme expliqué précédemment. Le substrat est ensuite contraint localement dans Plumed par un mur de potentiel (2 Å), permettant de limiter la distance interatomique entre le site actif de la protéine et l'un des atomes lourds du substrat. Aucune contrainte supplémentaire n'a été appliquée au système afin de permettre au substrat de s'adapter continuellement au site actif de la protéine et à la dynamique du solvant.

I. 5. Estimation de la sélectivité

hPRDX2 a été modélisée à partir de la structure cristallographique résolue à 1.7 Å (PDB : 1QMV).¹⁷⁰ L'homodimère a été construit à partir des chaînes A et B avec la modification du résidu N-carboxyalanine (résidu NCB) en alanine, et de la cystéine catalytique Cys50 suroxydée Cys-SO₂⁻ en cystéine classique Cys-S⁻. De plus une grande partie des atomes en dehors du site actif ont été contraints (100 kcal.mol⁻¹.Å⁻²): les Cα de la chaîne B (résidus 198 à 363) ainsi que ceux occupant l'extrémité *N-term* de la chaîne A (résidus 167 à 194), afin d'éviter une déstructuration du dimère au cours de la simulation (Figure 37).

I. 6. Clustering géométrique

Pour identifier les différentes conformations géométriques de liaison existantes lors d'une simulation, nous avons utilisé un module du logiciel de visualisation : VMD Clustering V 2.0 (Luis Garcia, Cornell Medical College, USA). La trajectoire est dans un premier temps chargée dans le logiciel VMD. Les « clusters » sont ensuite déterminés avec un cutoff de 2 Å en termes de RMSD, sur l'ensemble des cadres chargés.

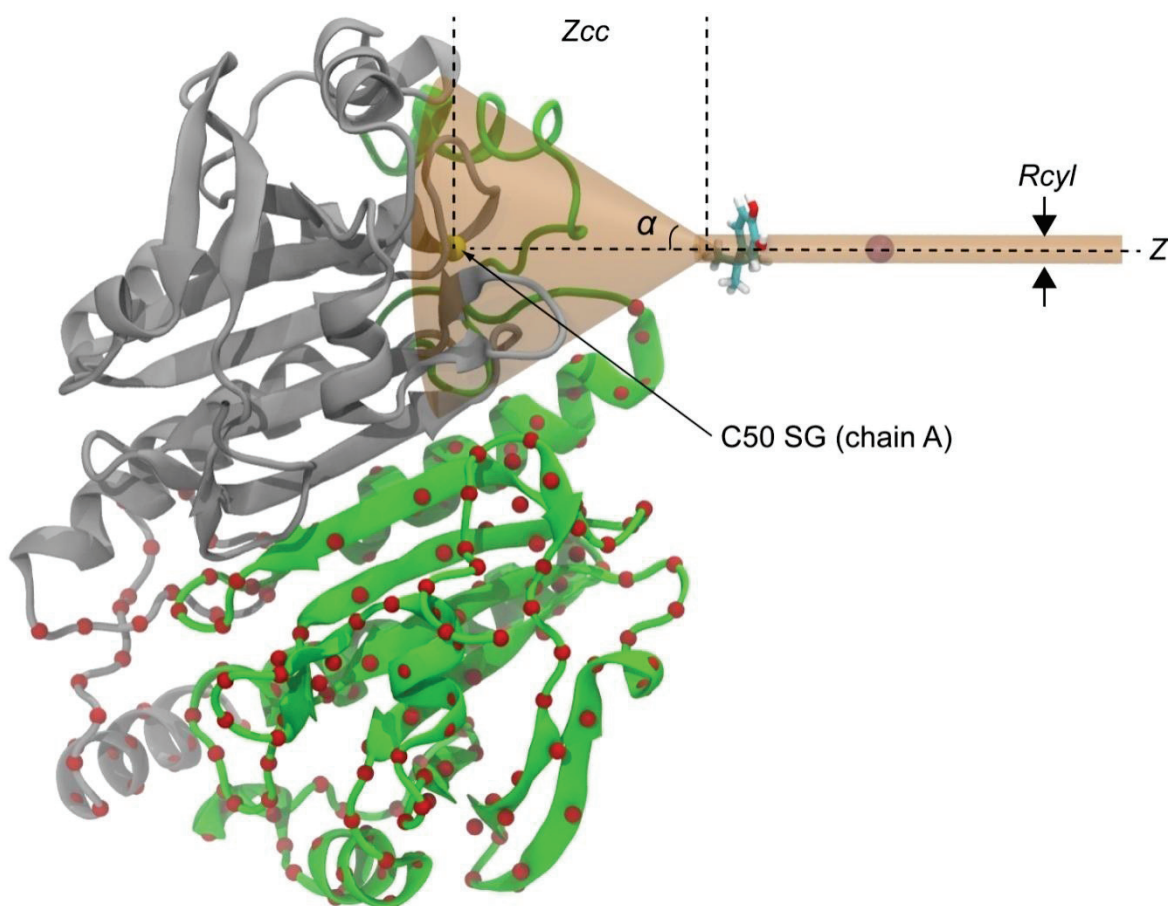


Figure 37. Structure de l'homodimère de hPRDX2 utilisé en FM.

Chaînes de l'homodimère colorées en gris (chaîne A) et vert (chaîne B). Visualisation des contraintes spatiales appliquées aux atomes lourds du système (sphères rouges), ainsi que de la position et de l'orientation du potentiel en forme d'entonnoir défini par l'axe Z, l'angle α , R_{cyl} le rayon de la partie cylindrique et Z_{cc} la distance entre l'origine de l'axe Z et le passage du cône au cylindre. L'axe Z a pour origine SG C50 et passe par un des atomes lourds du ligand (à définir à chaque simulation)

I. 7. Docking

Le *docking* a été effectué avec AutoDockVina 2.1.1 via le logiciel PyMol.⁹⁶ Ce programme de *docking* utilise les outils AutoDock Tools, Autogrid 4 et AutoDock 4.¹⁷¹ La grille utilisée ($40 \times 40 \times 40 \text{ \AA}^3$) a été centrée sur la cystéine catalytique de hPRDX5, et englobe le site actif ainsi que le ligand considéré. Les fichiers de structure de la protéine et du ligand sont utilisés après suppression des hydrogènes, et la protéine est modélisée avec ses chaînes latérales flexibles (excepté pour les acides aminés Pro, Ala et Gly). Les 10 meilleures poses obtenues sont visualisées via PyMol, et leur énergie libre de liaison est obtenue via la fonction de *scoring* empirique d'AutoDock qui repose sur la somme des contributions des interactions inter- et intramoléculaires du ligand.⁹⁶

II. PROTOCOLES DE PRODUCTION ET DE PURIFICATION DES PROTEINES

Les réactifs utilisés pour la réalisation des solutions tampons et des milieux de culture ont été obtenus commercialement par le biais des compagnies suivantes : Thermo Fisher Scientific (Fair Lawn, USA), Carl Roth GMBH + Co. KG (Karlsruhe, Germany), et Sigma-Aldrich (St. Louis, USA). Les colonnes utilisées en chromatographie viennent des compagnies Qiagen (Hilden, Germany) et GE Healthcare Life Sciences (Buckinghamshire, United Kingdom). La centrifugeuse utilisée pour la production et la purification des protéines est une Thermo Fisher Scientific possédant des rotors F15-8 x 50 cy et TX-750. Le sonicateur provient de la compagnie Sartorius (Sartorius Stedim Biotech, Aubagne, France) et le spectromètre UV-Visible de chez Jasco (Jasco France SAS, France).

II. 1. Expression des isoformes 1, 2 et 5 des peroxyrédoxines humaines recombinantes

Les protéines hPRDX1, hPRDX2 et hPRDX5 ont été exprimées sans leur séquence cible mitochondriale avec une étiquette hexahistidine (6xHis) en *N-term*, selon le protocole publié par Declercq et al.¹¹⁵ Des cellules M15 (pRep4) de *E.coli* ont été transformées en utilisant un plasmide pQE-30 (Qiagen) codant pour les protéines d'intérêt et conférant aux cellules transformées une résistance à deux antibiotiques : l'ampicilline et la kanamycine. Les cellules et le plasmide nous sont été généreusement fournies par le Pr. Knoop (Université de Louvain, Belgique). Les cellules transformées ont tout d'abord été déposées sur une gélose nutritive sur boîte de Pétri (milieu LB (Tableau 2) + agar) puis mise à incubées sur la nuit à 37°C en présence des deux antibiotiques précités.

Une colonie est ensuite prélevée et inoculée dans une pré-culture en milieu riche LB (Tableau 2) à 37°C contenant les 2 antibiotiques à 50 µg/mL chacun, sur la nuit. 25 à 30 mL de cette pré-culture sont ensuite prélevés et ajoutés pour 1 L de milieu de culture riche LB, et mis à incuber à 37°C à 150 rpm durant 2-3 h. L'expression de la protéine est ensuite induite par ajout d'isopropyl-thiogalactopyranoside (IPTG) (1 mM) à une $DO_{600nm} = 0,6 - 0,7$, pendant 4 heures à 37°C. Enfin, les cellules sont collectées par centrifugation (3000 rpm, rotor TX-750, 20min, 4°C) et stockées à -80°C.

II. 2. Expression de protéines marquées

Afin de produire des protéines marquées uniformément à l'azote ¹⁵N, les cellules bactériennes ont été incubées à 37°C à 150 rpm dans une pré-culture en milieu riche, puis dans une pré-culture en milieu minimum M9 (Tableau 2) contenant 1 g/L de ¹⁵NH₄Cl comme source de ¹⁵N, à

laquelle on a ajouté 1 % en volume du milieu de pré-culture riche. Elles ont ensuite été cultivées dans ce même milieu minimum M9 contenant du $^{15}\text{NH}_4\text{Cl}$, et l'expression a été induite de la même manière que pour la production de protéines non marquées. De la même manière, pour produire des protéines doublement marquées à l'azote ^{15}N et au carbone ^{13}C , les cellules ont été cultivées en milieu minimum M9 contenant cette fois 1 g/L de $^{15}\text{NH}_4\text{Cl}$ et 4 g/L de $^{13}\text{C}_6$ -D-glucose pour un marquage uniforme en $^{15}\text{N}/^{13}\text{C}$.

Tableau 2. Composition des milieux de culture

Milieux	Composition pour 1L de solution
Milieu riche LB (Luria Bertani)	25 g de LB (Extrait de levure, tryptone et NaCl) 1L dd H ₂ O
Milieu minimum M9¹⁷²	7.5 g de Na ₂ HPO ₄ ; 2H ₂ O 3 g de KH ₂ PO ₄ 0.5 g de NaCl 1 L dd H ₂ O
Suppléments au milieu M9	1 mM de MgSO ₄ 6 µg/L de thiamine 1 mM de CaCl ₂ 10 mL de solution traces de métaux* 20 mL de glucose filtré 20% (m/V)
Suppléments au milieu M9 pour marquage sélectif	1 g de $^{15}\text{NH}_4\text{Cl}$ pour une culture marquée ^{15}N 4 g de $^{13}\text{C}_6$ -D-glucose pour une culture marquée ^{13}C

* Solution de métaux traces composée de 5g d'EDTA, 0.5g de FeCl₃, 0.005g de ZnO, 0.001g de CuCl₂, 0.001g de CO(NO₃)₂.6H₂O et 0.001g de (NH₄)₆Mo₇O₂₄.4H₂O par litre de solution.

II. 3. Purification des isoformes 1, 2 et 5 des peroxyrédoxines humaines

Avant purification, les culots de cellules obtenus par centrifugation à 3000 rpm en fin de culture ont été resuspendus dans un tampon nommé tampon de lyse (Tableau 3) auquel est ajouté du lysozyme permettant de favoriser la lyse des cellules bactériennes. Ces cellules sont ensuite lysées par ultrasons durant 4 cycles de 3 min. Le surnageant de l'extrait cellulaire est ensuite récupéré par centrifugation (12000 rpm, rotor F15-8 x 50 cy, 45 min, 4°C). La purification des protéines a ensuite été réalisée sur une colonne d'affinité au nickel Ni²⁺-NTA (acide nitrilotriacétique) tirant partie de la

présence de l'étiquette histidine ajoutée en *N-term* de la protéine. La colonne est tout d'abord équilibrée avec le tampon de lyse avant d'y injecter le sumageant. La colonne est ensuite lavée avec le tampon de lyse pour éliminer toutes les espèces non retenues, puis la protéine est éluée en effectuant un gradient en imidazole grâce au tampon d'éluion concentré en imidazole (Tableau 3). Les fractions correspondant au pic de la protéine sont récupérées et regroupées.

La protéine ainsi récupérée est dialysée à 4°C durant 4 h contre du tampon phosphate salin (PBS) pH 7.4. La protéine purifiée est ensuite concentrée par centrifugation (6000 rpm, rotor F15-8 x 50 cy, 4°C) à l'aide de tubes Vivaspin (Sartorius) ayant un seuil de 10 kDa, et sa concentration déterminée par mesure d'absorbance à 280 nm connaissant le coefficient d'extinction molaire de hPRDX5 sous forme réduite déterminé à partir de la séquence en acides aminés de la protéine suivante : MRGSHHHHHH GSAPIKVGDA IPAVEVFEGE PGNKVNLAEL FKGKKGVLFG VPGAFTPGCS KTHLPGFVEQ AEALKAKGVQ VVACLSVNDA FVTGEWGRAH KAEGKVRLLA DPTGAFGKETD LLLDDSLVSI FGNRRLLKRF5 MVVQDGIVKA LNVEPDGTGL TCSLAPNIIS QL, $\epsilon = 5\,500 \text{ L}\cdot\text{mol}^{-1}\cdot\text{cm}^{-1}$ (expasy.org/protparam).

Tableau 3. Composition des solutions tampons utilisées

Tampons	Composition
Tampon PBS pH = 7.4 (Tampon de dialyse)	10 mM Na ₂ HPO ₄ 137 mM NaCl 3 mM KCl
Tampon de lyse pH = 8	50 mM NaH ₂ PO ₄ 300 mM NaCl 10 mM Imidazole
Tampon d'éluion pH = 8	50 mM NaH ₂ PO ₄ 300 mM NaCl 250 mM Imidazole

II. 4. Electrophorèse sur gel SDS

Afin de vérifier la pureté des protéines, des électrophorèses sur gel de polyacrylamide ont été effectués en présence de dodécylsulfate de sodium (SDS). Cette technique permet la séparation des protéines selon leur poids moléculaire (Figure 38), dans une cuve de tampon de migration (Tableau 4) traversée par un courant électrique. Les échantillons sont dans un premier temps

focalisés sur un gel de concentration par un courant de 90 V, puis séparées à 120 V sur un gel de séparation (compositions des gels données dans le Tableau 4).

Une fois la séparation terminée, le gel est coloré à l'aide de bleu de Coomassie, puis décoloré à l'aide d'un mélange acide acétique/éthanol/ eau (10% - 40% - 50%). Les bandes observées sont alors comparées avec celles obtenues pour le marqueur de poids moléculaire afin d'avoir la taille des protéines analysées (kDa). Si l'échantillon protéique purifié ne présente qu'une seule bande, la purification est optimale.

Les échantillons protéiques déposés sur gel ont été préparés selon le protocole suivant : 10 μ L de solution colorée au bleu de bromophénol concentrée x2 (Tableau 4) sont ajoutés à 10 μ L de solution de protéine. Le mélange est ensuite chauffé durant 5 min au bain marie à 95°C.

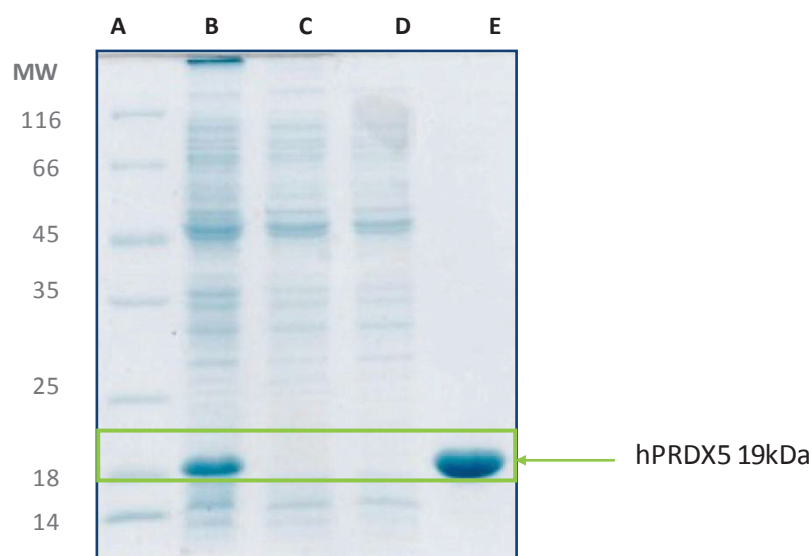


Figure 38. Exemple de gel SDS obtenu par électrophorèse en conditions dénaturantes pour la production de hPRDX5 non marquée.

A : marqueur de poids moléculaire, B : surnageant de l'extrait cellulaire, C : sortie de colonne lors du passage du surnageant, D : sortie de colonne lors du lavage, E : protéine récupérée après élution et concentration. L'échantillon protéique purifié ne présente qu'une bande intense correspondant à la protéine d'intérêt. La protéine obtenue est donc pure, et aucune perte n'est observée lors de la purification (aucune trace en sortie de colonne).

Tableau 4. Composition des gels et tampons utilisés en électrophorèse

- Composition des gels

Produit	Gel de concentration	Gel de séparation
Pourcentage	5%	15%
Polyacrylamide 30%	850 uL	7.5 mL
Tris 1M	625 uL (pH = 6.8)	3.75 mL (pH = 8.8)
H₂O	3.4 mL	3.5 mL
SDS 10% (m/V)	50 µL	150 µL
TEMED	5 µL	6 µL
Ammonium persulfate 10%	50 µL	150 µL
Volume total	≈ 5 mL	≈ 15 mL

- Composition du tampon de migration

Produit	Quantité	Concentration
Tris base	3 g	0.025 M – pH8.3
Glycine	14.4 g	0.19 M
SDS 10% (m/V)	10 mL	0.1 %
H₂O	990 mL	-

- Composition du tampon au bleu de bromophénol (V = 10 mL) concentré deux fois

Produit	Volume	Concentration
Tris - HCl 0.5 M – pH 6.8	2.5 mL	62.5 mM
SDS 10%	2.0 mL	1 %
Glycérol 100 %	3.0 mL	5 %
β-mercapto-ethanol 14.4 M	0.2 mL	150 mM
Bleu de bromophénol	Pointe de spatule	-
H₂O	2.3 mL	-

II. 5. Expression et purification de la thioredoxine (Trx) et de la thioredoxine réductase (TrxR)

Afin de pouvoir étudier la cinétique de réaction de hPRDX5, les autres protéines participant à son cycle catalytique: la thiorédoxine (γ Trx) et la thiorédoxine réductase (γ TrxR) toutes deux de levure, ont été produites en absence de marquage isotopique à partir de plasmides pQE-30 codant pour ces deux protéines et fournis par le Pr. Knoops. L'expression de ces deux protéines a été effectuée de manière identique à celle de hPRDX5. Ces deux protéines disposant également d'une étiquette histidine, elles ont été purifiées identiquement à hPRDX5 (voir II.3 p.88).

III. SYNTHÈSES D'INHIBITEURS

Les molécules fragments obtenues commercialement (Sigma-Aldrich) ont été utilisées sans étape de purification supplémentaire. La synthèse des différents inhibiteurs non commerciaux étudiés a été effectuée dans les locaux de l'Institut de Chimie et de Biochimie Moléculaires et Supramoléculaires (ICBMS), dans l'équipe de Chimie Organique 2 – Glycochimie, en collaboration avec le Dr. Sebastien VIDAL.

Les réactions ont toutes été réalisées sous atmosphère inerte (Argon (Ar)). Le suivi des réactions à été réalisé par chromatographies sur couche mince (CCM) sur des feuilles d'aluminium couverte d'une fine couche de silice 60 F254 (Merck) puis révélées par UV à 254 nm. La purification des produits a été effectuée sur des colonnes remplies de gel de silice Si 60 (40-63 μm). Les spectres RMN d'identification et de caractérisation des différents intermédiaires réactionnels ainsi que des produits finaux ont été enregistrés à 303K sur un spectromètre Bruker Avance III HD 400 MHz. La valeur du déplacement chimique δ (ppm) est référencée par rapport aux pics résiduels du solvant deutéré utilisé (majoritairement le chloroforme CDCl_3 ou le DMSO- d_6).

Les multiplicités des différents pics observés sont décrites en utilisant les abréviations suivantes : s, singulet; d, doublet; t, triplet; q, quadruplet; m, multiplet (Annexe 3).

III. 1. Synthèse du 4-néopentylbenzene-1,2-diol

J'ai synthétisé le 4-neopentylbenzene-1,2-diol (4-néopentylcatéchol (4NPC)) **4** en 3 étapes à partir du 1,2-diméthoxybenzene **1** (Figure 39).

L'acylation a été effectuée dans un liquide ionique comme décrit par Tran *et al.* avec $\text{Bi}(\text{OTf})_3$ comme catalyseur.¹⁷³ La cétone aromatique **2** a ensuite été sélectivement réduite en se basant sur le mode opératoire publié par l'équipe d'E. Metay à l'ICBMS, qui utilise un solvant vert (cyclopentyl methyl ether – CPME) avec du palladium sur charbon (Pd/C) 5% come catalyseur.¹⁷⁴ Enfin les éthers méthyliques ont été réduits à l'aide de BBr_3 .

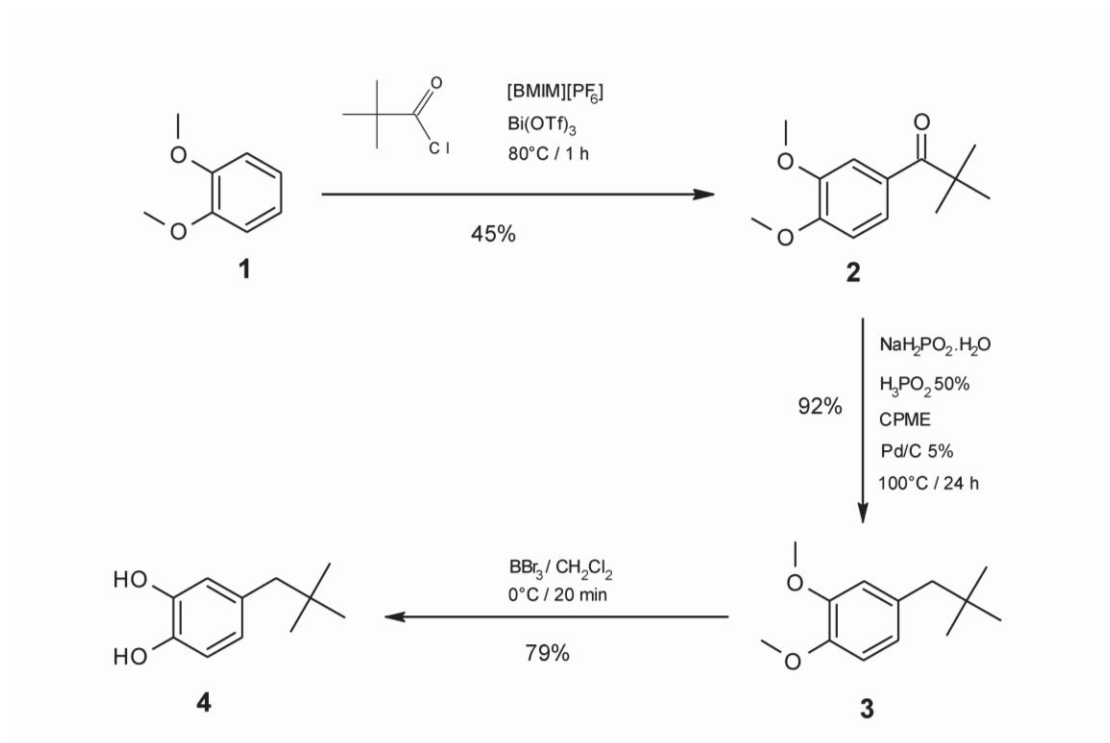


Figure 39. Schéma réactionnel d'obtention du 4NPC [4] à partir du 1,2-diméthoxybenzène [1].
Avec $[\text{BMIM}][\text{PF}_6]$: 1-butyl-3-méthylimidazoliumhexafluorophosphate. Me = CH_3

III. 1. 1 1-(3,4-diméthoxyphényl)-2,2-diméthylpropan-1-one [30314-46-6]

Le 1,2-diméthoxybenzène **1** (10 mmol, 1.27 mL), le $\text{Bi}(\text{OTf})_3$ (0.5 mmol, 328 mg, 0.5 éq) et le chlorure de pivaloyl (10 mmol, 1.23 mL, 1 éq) ont été agités à 80°C dans le liquide ionique $[\text{BMIM}][\text{PF}_6]$ (5 g) durant 1 h jusqu'à la disparition totale du réactif de départ. Le suivi de la réaction a été effectué par CCM (CH_2Cl_2 /Ether de pétrole (EP), 3:2). Après refroidissement, le milieu réactionnel a été extrait avec de l'éther diéthylique (4x50 mL) par sonication pour améliorer la séparation de la phase organique et du liquide ionique. Les phases étherées ont été regroupées et lavées avec une solution saturée de NaCl (4x30 mL), séchées sur Na_2SO_4 et concentrées. Le produit résiduel brut a été purifié par flash chromatographie (CH_2Cl_2 /EP, 1:1 à 4:1) pour donner le 1-(3,4-diméthoxyphényl)-2,2-diméthylpropan-1-one **2** (1 g, 45%) pur sous forme d'huile jaune.

III. 1. 2 4-(2,2-diméthylpropyl)-1,2-diméthoxybenzène [3]

Une solution d'hypophosphite de sodium monohydraté (7.5 mmol, 795 mg, 3 éq) et d'acide hypophosphorique H_3PO_2 à 50% (2.5 mmol, 165 mg, 1 éq) dans de l'eau permutée (5 mL) a été ajoutée à un mélange contenant la cétone **2** (2.5 mmol, 555 mg) et le catalyseur Pd/C 5% (0.25 mmol, 530 mg, 0.1 éq) dans du CPME (2.5 mL). Le milieu réactionnel a été chauffé à 100°C sous atmosphère inerte (Ar) pendant 24 h.

La réaction a ensuite été stoppée par l'addition de CH_2Cl_2 (30 mL) et d'eau (40 mL). La phase aqueuse a été extraite avec du CH_2Cl_2 (4x75mL). Les phases organiques ont été combinées, séchées sur Na_2SO_4 , filtrées sur célite et lavées avec CH_2Cl_2 (100 mL). Le solvant a ensuite été évaporé pour donner le 4-(2,2-diméthylpropyl)-1,2-diméthoxybenzène **3** (448 mg, 92%) sous forme de liquide jaune clair sans purification. Le produit a cristallisé à 4°C après quelques heures, pour donner un solide de couleur jaune pâle.

III. 1. 3 4-(2,2-diméthylpropyl)benzène-1,2-diol [4]

Une solution de BBr_3 (4.5 mmol, 4.5 mL, 1M dans CH_2Cl_2 , 3.2 éq) a été ajoutée sous atmosphère inerte, et à 0°C, à une solution du composé **3** (1.4 mmol, 290 mg) dans du CH_2Cl_2 (15 mL). Le milieu réactionnel de couleur rouge brique a ensuite été agité à 0°C pendant 20 min. La réaction a été stoppée rapidement par addition d'une solution aqueuse saturée de carbonate de sodium NaHCO_3 (15 mL) avant d'être agitée vivement durant 2,5 h. Après dilution du milieu réactionnel dans CH_2Cl_2 (60 mL), la phase organique a été extraite avec une solution aqueuse de thiosulfate de sodium $\text{Na}_2\text{S}_2\text{O}_3$ à 10% (50 mL) pour éliminer le brome résiduel, puis lavée à l'eau glacée (30 mL) avant d'être séchée sur Na_2SO_4 et filtrée sur célite. Le solvant a été évaporé pour obtenir le produit **4** brut sous forme de poudre blanche cristalline. Une purification par flash chromatographie sur gel de silice (éluant : $\text{CH}_2\text{Cl}_2/\text{CH}_3\text{OH}$, 98:2) a été effectuée afin d'obtenir des échantillons de pureté optimale pour les essais biologiques, et a permis d'obtenir le 4-(2,2-diméthylpropyl)benzène-1,2-diol **4** (198 mg, 79%) sous forme de poudre blanche cristalline.

III. 2. Synthèse du 4-néopentyl-5-phenylaminobenzène-1,2-diol

Le 4-neopentyl-5-phenylaminocatechol **6** a été synthétisé par Yoann Pascal, doctorant à l'ICBMS de Lyon, en 4 étapes à partir du 1-(3,4-diméthoxyphenyl)-2,2-diméthylpropan-1-one **2** obtenu lors de la synthèse du 4-néopentylcatéchol (Figure 40).

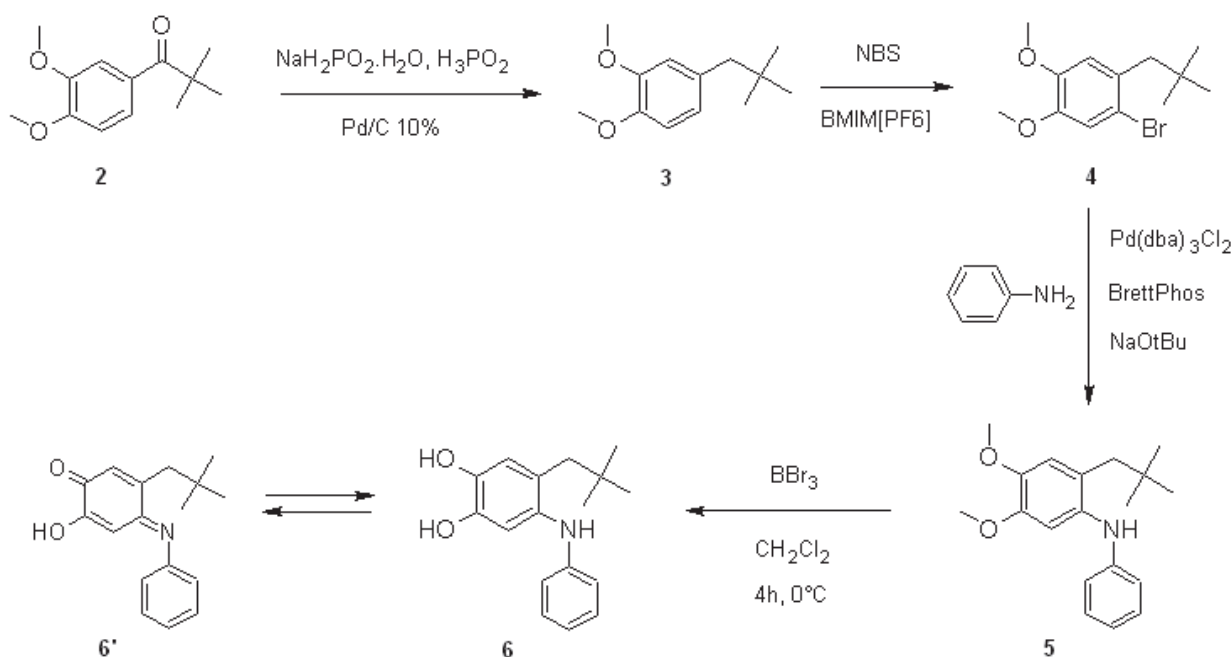


Figure 40. Schéma réactionnel d'obtention des ligands [6] et [6'] à partir de [2].

Avec NBS : *N*-bromosuccinimide et Brett Phos un ligand phosphine utilisé dans la réaction de Buchwald.

III. 2. 1 4-(2,2-diméthylpropyl)-1,2-diméthoxybenzène [3]

La cétone **2** a été introduite (2.34 mmol, 520 mg) dans un tube de micro-onde et dissoute dans 2.5 mL de CPME. L'hypophosphite de sodium monohydraté (7.02 mmol, 744 mg, 3 éq) et le Pd/C 10% (0.234 mmol, 249 mg, 0.1 éq) ont ensuite été ajoutés ainsi que 5 mL d'eau distillée. Enfin, l'acide hypophosphoreux 50% m/V (1.17 mmol, 80 mg, 0.5 éq) a été introduit. Le tube a été scellé, et chauffé à 100°C au micro-onde pendant 24 h. En fin de réaction, la solution a été diluée avec du CH₂Cl₂ (20 mL) et de l'eau distillée (40 mL). La phase aqueuse a été extraite avec du CH₂Cl₂ (4x75 mL). Les phases organiques ont été combinées, séchées sur MgSO₄ puis filtrées sur célite et lavées avec du CH₂Cl₂ (150 mL). Enfin, le solvant a été évaporé pour donner le produit **3** (448 mg, 92%) sous la forme d'une huile légèrement jaune.

III. 2. 2 1-bromo-2-(2,2-diméthylpropyl)-4,5-diméthoxybenzène [4]

Le composé **3** (2.15 mmol, 448 mg) a été placé dans du liquide ionique [BMIM][PF₆] (7g). Le système a été placé sous atmosphère inerte puis le *N*-bromosuccinimide (NBS) (2.26 mmol, 402 mg, 1.05 éq) y a été ajouté. La réaction a été chauffée à 60°C pendant 2 h puis agitée à température ambiante durant 15 h. Le milieu réactionnel a été extrait avec de l'éther diéthylique (3 x 20 mL) et du

toluène (20 mL) par sonication pour améliorer la séparation de la phase organique et du liquide ionique. Les phases contenant l'éther et le toluène ont été réunies et évaporées. Le produit brut a été purifié sur colonne de silice (EP/acétate d'éthyle, 8:2) donnant un mélange du 1-bromo-2-(2,2-diméthylpropyl)-4,5-diméthoxybenzène **4** et du produit de départ (520 mg, 84% dont 7% de produit de départ) sous la forme d'une huile légèrement jaune.

III. 2. 3 2-(2,2-diméthylpropyl)-4,5-diméthoxy-N-phénylaniline [5]

Le *tert*-butylate de sodium NaOtBu (3.62 mmol, 347 mg, 2 éq), le tris (dibenzylidèneacétone) dipalladium dichlore Pd₂(dba)₃Cl₂ (0.037 mmol, 39 mg, 0.02 éq) et la BrettPhos (0.113 mmol, 39 mg, 0.06 éq) ont été introduits dans un Schlenck, et le système a été placé sous atmosphère inerte (Argon). **4** (1.81 mmol, 520 mg) a été dissout dans du *tert*-butanol tBuOH (20 mL) préalablement dégazé durant 2 h (technique de bullage par Ar). Ce mélange a alors été introduit dans le milieu réactionnel ainsi que l'aniline (3.62 mmol, 330 µL, 2 éq). La réaction a été chauffée à 100°C durant 16 h dans un bain d'huile déjà préchauffé. Le solvant a été évaporé et le produit brut récupéré a alors été purifié sur colonne de silice (EP/acétate d'éthyle, 9:1) pour donner le 2-(2,2-diméthylpropyl)-4,5-diméthoxy-N-phénylaniline **5** (415 mg, 77%) sous la forme d'un solide légèrement marron.

III. 2. 4 4-(2,2-diméthylpropyl)-5-(phénylamino)benzène-1,2-diol [6]

Le composé **6** (0.334 mmol, 100 mg) a été dissout dans CH₂Cl₂ (5 mL). Le milieu a été placé sous atmosphère inerte (Ar) et refroidi à 0°C à l'aide d'un bain de glace. BBr₃ 1M dans CH₂Cl₂ (1 mmol, 1 mL, 3 éq) a ensuite été ajouté goutte à goutte. Une fois revenu à température ambiante, le système est agité pendant 4 h. NaOH 1 M (5 mL) a été ajouté et la réaction a été laissée sous agitation durant 22 h. HCl 2 M a alors ajouté jusqu'à atteindre un pH = 1. Le mélange réactionnel a été dilué avec de l'eau (20 mL) et la phase aqueuse extraite avec CH₂Cl₂ (3 × 40 mL). Les phases organiques ont été réunies, séchées sur MgSO₄, filtrées et évaporées. Le 4-(2,2-diméthylpropyl)-5-(phénylamino)benzène-1,2-diol **6** (60 mg, 67%) a alors été récupéré sous sa forme oxydée **6'** (≈ 90 %) en équilibre avec sa forme réduite (≈ 10 %), avec un aspect de solide de couleur rouge brique.

IV. SPECTROSCOPIE

La totalité des expériences de RMN en solution ont été réalisées au laboratoire, la majorité d'entre elles sur un spectromètre Varian Inova 600MHz équipé d'une sonde standard à champs de gradient pulsés (PFG) 5mm triple résonance ($^1\text{H}/^{13}\text{C}/^{15}\text{N}$). Les dernières expériences, ont été menées sur ce même spectromètre à présent équipé d'une sonde cryogénique 5mm triple résonance PFG « $^1\text{H}/^{13}\text{C}$ enhanced ». Enfin, les expériences de caractérisation de molécules organiques ont été réalisées sur un spectromètre Bruker Avance II HD 400MHz équipé d'une sonde 5 mm multinoyaux PABBO BB/19F – 1H/D Z-GRAD. La transformation et l'analyse des spectres ont été faites à l'aide des logiciels NMRPipe, NMRDraw,¹⁷⁵ et NMRViewJ¹⁷⁶ pour les expériences effectuées sur le spectromètre Varian, et des logiciels Bruker TopSpin 3.2 et NMRNotebook 2.70 pour les expériences effectuées sur le spectromètre Bruker et les expériences à 1 dimension.

IV. 1. Caractérisation de molécules organiques par RMN et SM

Les expériences RMN utilisées pour la caractérisation des intermédiaires et des produits finaux de synthèse ainsi que l'attribution complète des molécules synthétisées a été possible grâce à l'acquisition de spectres 1D (^1H et ^{13}C) et 2D (COSY, HSQC et HMBC si nécessaire). Les paramètres utilisés pour ces expériences sont compilés dans le Tableau 5.

De plus, des expériences de spectrométrie de masse haute résolution (HR-ESI-QToF) ont été effectuées sur un spectromètre Bruker MicroToF-Q II XL. La spectrométrie de masse haute résolution utilisée est une expérience en tandem HR-ESI-QToF qui couple une ionisation par électronébuliseur à un quadripôle relié à un analyseur à temps de vol.

IV. 2. Expériences 2D ^{15}N HSQC

Pour les expériences HSQC, les tubes RMN ont été préparés dans un tampon PBS de pH = 7.4 avec usuellement 200 μM de protéine marquée uniformément en ^{15}N et des ajouts de concentration croissante de fragments solubilisés dans du diméthylsulfoxyde deutéré (DMSO- d_6). Dans le cas du ligand 4NPC, la concentration de la protéine était de 120 μM . La concentration finale de DMSO- d_6 dans les tubes RMN n'excède pas 8%. La forme réduite de la protéine hPRDX5 a été maintenue par la présence de 2 mM de TCEP. Les expériences ^{15}N -HSQC ont été enregistrées à 28°C avec les paramètres expérimentaux donnés dans le Tableau 5. Chaque expérience HSQC a été précédée d'une expérience 1D- ^1H pour vérifier la pureté et la stabilité de la protéine et des ligands.

Tableau 5. Expériences RMN et paramètres d'acquisition utilisés

Echantillon	Fréquence de travail (MHz)*	Expériences	Nombre de scans	Largeur spectrale (ppm)	Nombre de points
hPRDX5	600	1D - ^1H	512	15	16384
		1D - STD	1024	15	16384
		$^1\text{H}/^{15}\text{N}$ - HSQC	16	15 x 32	2048x64
Ligands	400	1D - ^1H	32	20	16384
		1D - ^{13}C	4096	239	16384
		$^1\text{H}/^1\text{H}$ - COSY	4	10 x 10	2048x256
		$^1\text{H}/^{13}\text{C}$ - HSQC	16	16 x 220	2048x512
		$^1\text{H}/^{13}\text{C}$ - HMBC	8	20 x 239	2048x512

* Fréquence de travail = fréquence de résonance du proton

➔ Caractéristiques générales pour le spectromètre Bruker Avance II HD 400 MHz :

Proton	Pulse 90 = 13.70 μs Relaxation delay = 2.0 s
Carbone	Pulse 90 = 9.25 μs Relaxation delay = 3.0 s

➔ Caractéristiques générales pour le spectromètre Varian Inova 600 MHz :

Sonde chaude standard 5mm triple résonance ($^1\text{H}/^{13}\text{C}/^{15}\text{N}$) PFG

Proton Pulse 90 = 7.2 μs
Relaxation delay = 2.0 s

Sonde cryogénique 5mm triple résonance PFG « $^1\text{H}/^{13}\text{C}$ enhanced »

Proton Pulse 90 = 8.1 μs
Relaxation delay = 1.0 s

IV. 3. Expériences STD

Les expériences par différence de transfert de saturation (STD) réalisées sur les différents complexes hPRDX5-ligand ont été effectuées à 28°C et les paramètres expérimentaux utilisés sont répertoriés dans le Tableau 5. Les tubes ont été préparés dans un tampon PBS (pH = 7.4) avec 10 ou 20 µM de protéine en présence de 500 µM de TCEP pour garantir son état réduit, et 10% de D₂O (v/v) pour la stabilité du champ magnétique (signal de lock). L'interaction des différents ligands a été étudiée en ajoutant des concentrations croissantes en ligands, en fonction de leur affinité déterminée par les titrations en HSQC. Pour chaque concentration en ligand, une expérience 1D-¹H et une expérience 1D-STD ont été enregistrées dans les mêmes conditions expérimentales, mais avec un nombre de scans deux fois plus important dans le cas de l'expérience STD. L'irradiation sélective de la protéine a été effectuée à - 0.23 ppm (voir spectre 1D ¹H de hPRDX5 en Annexe 1), située à - 3006 Hz de la fréquence de la porteuse (H₂O : 4.78 ppm) et un temps de saturation totale de 2 s composé de 40 impulsions de 50 ms séparées par un délai de 100 µs. Pour chaque concentration en ligand, le facteur STD (f_{STD}) a été calculé (équation 35), ce qui permet de tracer la courbe $f_{STD} = f([L])$ en fin d'expérience. Les résultats ont ensuite été normalisés par rapport au proton ayant le plus grand f_{STD} (100%) afin de réaliser l'*epitope mapping*.

$$f_{STD} = \frac{I_{STD}}{I_{1D}} \cdot \frac{[ligand]}{[protéine]} \quad [35]$$

Avec f_{STD} le facteur STD, I_{STD} la valeur de l'intégrale du signal concerné sur le spectre STD irradié sélectivement, I_{1D} la valeur de l'intégrale de ce même signal sur le spectre 1D ¹H sans irradiation sélective, et $[L]/[P]$ le rapport des concentrations.

Les points expérimentaux sont ensuite ajustés avec une courbe théorique de type hyperbole :

$$y = \frac{Ax}{B+x} \quad [36]$$

Avec A : la valeur maximale de saturation (plateau supérieur) ; et B : la valeur du K_D (même unité que la concentration).

IV. 4. Mesure de l'affinité protéine-ligand par RMN

Les perturbations de déplacement chimique (CSPs) ont été calculées selon l'équation 37 :

$$CSP = \sqrt{\Delta\delta_H^2 + \frac{1}{25} \Delta\delta_N^2} \quad [37]$$

Avec CSP la variation de déplacement chimique moyenne induite par l'interaction du ligand sur un signal donné, $\Delta\delta_H$ sa valeur sur l'axe du proton ¹H et $\Delta\delta_N$ sa valeur sur l'axe de l'azote ¹⁵N. En utilisant

un seuil de quantité de CSP arbitraire, défini pour chaque ligand, les acides aminés les plus impliqués dans l'interaction protéine-ligand sont caractérisés et utilisés pour le calcul de la constante de dissociation (K_D) du complexe.

Les courbes isothermes $CSP = f([L])$ sont tracées pour ces acides aminés, puis le K_D est ensuite extrapolé par corrélation des points expérimentaux de ces isothermes avec une régression non-linéaire via le programme MatLab (The Math Works, Inc) utilisant l'expression **38** suivante pour un modèle de liaison 1:1 impliquant qu'une seule molécule de ligand se lie à un seul site actif (voir la démonstration de l'obtention de cette équation en annexe 2):

$$\Delta\delta_{obs} = \Delta\delta_{max} \cdot \frac{([P]_0 + [L]_0 + K_D) - \sqrt{([P]_0 + [L]_0 + K_D)^2 - 4[P]_0[L]_0}}{2[P]_0} \quad [38]$$

Avec $\Delta\delta_{obs}$ la CSP observée expérimentalement pour chaque point de la titration, $\Delta\delta_{max}$ la CSP maximale observée (généralement lorsque la saturation de la protéine est atteinte, ou bien sa valeur au dernier point de la titration), $[P]_0$ et $[L]_0$ les concentrations initiales en ligand et en protéine, et K_D la constante de dissociation du complexe.^{177; 178}

V. EXPERIENCES DE CINETIQUE ENZYMATIQUE

V. 1. Détermination expérimentale de K_M et V_{max}

La détermination du K_M et du V_{max} pour les protéines hPRDX5, hPRDX1 et hPRDX2 a été effectuée à température ambiante ($\approx 23^\circ\text{C}$) en présence d'une concentration croissante de substrat H_2O_2 entre 0 et 500 μM . La vitesse initiale de réaction V_0 a été calculée indirectement par mesure de l'absorbance du NADPH à 340 nm (absorbance décroissante au cours du temps due à l'oxydation du NADPH par la TrxR (Figure 29)). Par représentation graphique des V_0 en fonction de la concentration en substrat, on obtient des courbes de type hyperbole où l'on retrouve la valeur de $[S] = K_M$ pour $\frac{1}{2}$ de V_{max} . Cette expérience permet aussi de pouvoir vérifier l'activité des protéines produites, pour le système protéique utilisé déterminé par M. Chow lors de sa thèse au laboratoire, composé de :¹

- 15 μM de γTrx ,
- 2 μM de γTrxR ,
- 0.15 μM de hPRDX5 (0.38 μM dans le cas des isoformes hPRDX1 et hPRDX2)
- 200 μM de NADPH
- 25 μM de H_2O_2

V. 2. Evaluation de l'IC₅₀

Le protocole de détermination de la concentration inhibitrice médiane a été adapté de Kim et al. 2005.179 Les réactions de détermination d'IC₅₀ dans le cas de hPRDX5 ont été menées sur 1000 μL , en présence de 15 μM de γTrx , 2 μM de γTrxR , 0.15 μM de hPRDX5, 200 μM de NADPH et 25 μM de H_2O_2 ($\approx K_M$), pour des concentrations en ligands (solubilisés dans le DMSO) variables selon leur affinité : catechol 0-50 mM, 4-tert-butylcatechol 0-5 mM, 4-neo-pentylcatechol 0-1.5 mM. Les protéines ont d'abord été diluées dans une solution de PBS (Tableau 3) comportant 1 mM d'EDTA à pH 7.0. Une fois ce mélange introduit dans la cuve UV, le ligand est ajouté (quantité maximale de DMSO ajoutée $\leq 5\%$), puis le NADPH, et enfin H_2O_2 qui va initier la réaction. La réaction a été suivie par UV à température ambiante à 340 nm durant 200 s car comme pour les mesures de K_M , l'activité de hPRDX5 a été indirectement mesurée par la mesure de l'absorbance du NADPH à 340 nm. La vitesse initiale V_0 de la réaction a été déterminée à partir d'une portion linéaire des courbes de vitesses enregistrées (environ 10 premières secondes), et exprimée en $\mu\text{mol}\cdot\text{min}^{-1}\cdot\text{L}^{-1}$. Les IC₅₀ relatifs ont ensuite été déterminés à la moitié du ΔV_0 en traçant les valeurs de V_0 en fonction de la concentration en ligand utilisée, et en faisant correspondre les points expérimentaux obtenus avec

une équation de type hyperbole (Equation 39) sur OriginPro 8 (OriginLab, Northampton, MA) et SigmaPlot 11.0 (Systat Software, San Jose, CA).

$$y = C + \frac{B \cdot A}{x + A} \quad [39]$$

Avec A : la valeur de l' IC_{50} relatif ; B : la valeur de l'asymptote verticale (en x) ; et C : la valeur de la vitesse résiduelle (en %).

Chaque expérience a été répétée et les valeurs d' IC_{50} obtenues sont des moyennes calculées sur 3 à 5 expériences selon les ligands.

Les IC_{50} relatifs des dérivés catéchols pour l'inhibition des isoformes hPRDX1 et hPRDX2 ont été mesurés dans des conditions expérimentales similaires, avec une quantité de protéine adaptée à 0.38 μ M.

V. 3. Etude du mécanisme d'inhibition

Pour l'étude du mécanisme d'inhibition de hPRDX5, 5 concentrations en ligands ont été utilisées dont 3 sont voisines de l' IC_{50} à l'erreur près (gamme de concentration allant de 0 à 50 mM pour 4NPC). Le reste du système reste inchangé avec 15 μ M de γ Trx, 2 μ M de γ TrxR, 0.15 μ M de hPRDX5, 200 μ M de NADPH et une gamme de H_2O_2 allant de 0 à 500 μ M. Comme pour l'expérience de détermination de l' IC_{50} , après mélange des trois protéines dans la solution de PBS contenant 1mM d'EDTA préparée à pH = 7.0, le dérivé catéchol est ajouté au mélange, puis le NADPH et enfin H_2O_2 qui va initier la réaction. L'intérieur de la cuve est homogénéisé après ajout du ligand préalablement solubilisé dans le DMSO, et après ajout de H_2O_2 , avant de démarrer l'enregistrement de l'absorbance à 340 nm durant 100s. Les expériences ont été traitées de la même manière que pour la mesure de l' IC_{50} , et les vitesses exprimées en μ mol.min⁻¹.L⁻¹. Les données ont ensuite été analysées avec le logiciel DynaFit en utilisant différents modèles de mécanisme d'inhibition (voir chapitre I, partie III.2.2).^{180,181} Le mécanisme d'inhibition se rapprochant le plus des points expérimentaux a été sélectionné. Les courbes correspondantes ont été tracées sur Origin 8.0 avec comme erreur, l'erreur standard sur chaque mesure.

CHAPITRE III

RESULTATS ET DISCUSSION

Dans ce chapitre, les résultats de ce projet seront présentés selon trois parties :

Dans un premier temps, la corrélation expérimentale et simulée des mesures d'affinités obtenues par RMN et FM sera démontrée. Pour cela nous nous sommes appuyés sur deux ligands dont l'affinité pour hPRDX5 a déjà été mise en évidence : le catéchol (CAT) et le 4-méthylcatéchol (4MEC).^{91; 182} Nous montrerons ici que les résultats obtenus par FM et RMN sont en corrélation. De plus, nous verrons que l'analyse des simulations FM donne de nombreuses informations statistiques sur l'équilibre d'interaction entre protéine et ligand.

Dans une deuxième partie, nous utiliserons les informations provenant des calculs de FM et nous présenterons l'étude de l'interaction de plusieurs autres homologues du catéchol avec hPRDX5. La détermination de l'affinité de ces ligands nous a permis d'optimiser pas à pas l'interaction protéine-ligand. Dans un projet de conception de molécules bioactives, nous montrerons que la FM permet l'obtention de nombreuses informations prédictives permettant l'optimisation de ligands d'affinité moyenne (millimolaire) en ligands de meilleure affinité (micromolaire).

hPRDX5 étant une enzyme, nous prouverons dans la troisième partie grâce à des expériences de cinétique enzymatique, que les résultats d'affinité obtenus par le calcul et prouvés par RMN sont proportionnels à l'activité inhibitrice des catéchols. L'étude du mécanisme d'inhibition des dérivés catéchols sera présentée et discutée en rapport avec la FM, ainsi que leur inhibition pour deux autres isoformes des hPRDX (hPRDX1 et hPRDX2) afin d'estimer leur sélectivité.

Nous pourrions alors conclure que la FM peut être utilisée afin de prédire quantitativement des affinités protéine-ligand, mais aussi qu'elle pourrait être utilisée comme outil dans la prédiction des mécanismes d'inhibition enzymatique induits par les ligands.

I. LA FUNNEL-METADYNAMIQUE ET LA RMN EN SOLUTION POUR L'ESTIMATION DE L'AFFINITE PROTEINE-LIGAND

Dans le chapitre d'introduction, nous avons présenté une technique particulière de dynamique moléculaire, la funnel-métadynamique (FM). Pour prouver que les résultats obtenus par dynamique estimaient de façon précise l'affinité d'un complexe PL, nous nous sommes appuyés sur l'étude du catechol (CAT), et de son dérivé méthylé : le 4-méthylcatechol (4MEC). Ces résultats ont été publiés en 2015 dans le *Journal of the American Chemical Society*.¹⁸³

I. 1. Le protocole de FM pour l'étude du système hPRDX5

I. 1. 1 Préparation du système

- Structure initiale

Dans un projet d'étude par MD, la structure de départ est un pré-requis essentiel. Comme présenté en chapitre d'introduction (voir III.1.1 p.51), plusieurs structures cristallographiques de hPRDX5 sont disponibles. En 2013, la structure possédant la meilleure résolution cristallographique (1.45 Å) était celle déposée par Hall et al. en 2010 (3MNG.pdb), qui possède également un ligand dans son site actif (DTT oxydé).¹⁰⁶ Des molécules d'eau, des ions (Na^+ , Cl^- et Br^-), ainsi que du glycérol sont aussi définis dans le cristal. Hall et al. discutent la présence de ligands dans les différentes structures cristallographiques de hPRDX5 (dont le DTT oxydé dans 3MNG.pdb) comme pouvant être des analogues de l'état de transition de l'enzyme. On sait que les enzymes dont le rôle fondamental est de stabiliser l'état de transition, se lient en général fortement à un ligand analogue de l'état de transition (voir III.3.4 du chapitre I p.72).¹⁵⁵

Pour la préparation de la MD, il y a plusieurs options possibles à partir de cette structure cristallographique détaillée. Il est possible de ne conserver que la protéine et d'éliminer les sels, ligands et molécules d'eau en éditant le fichier PDB. L'eau et les ions seront par la suite remplacés lors des étapes de solvation et de définition de la salinité du système. Toutefois, les protéines possèdent parfois des cavités hydratées par quelques molécules d'eau seulement. Or il a été montré par RMN dans les années 90 que ces molécules d'eau avaient une cinétique d'échange très lente avec le solvant.¹⁸⁴ Dans le cas de hPRDX5 nous ne disposons pas de cette information, mais des cavités de ce type existent. Si nous supprimions toutes les molécules d'eau du cristal nous avons un risque de créer un vide difficile à réhydrater lors de l'équilibration du système malgré l'étape de solvation préalable de la protéine. Nous avons donc choisi par prudence de conserver les

molécules d'eau définies dans le cristal pour l'étape de solvation. Néanmoins, des systèmes ont aussi été modélisés en parallèle sans ces molécules d'eau, sans grande différence visible. Cela reste une question à se poser à l'origine d'un projet de MD, et importante pour la définition du système macromoléculaire étudié.

Un autre point important dans la préparation initiale d'un système pour la MD est l'état d'ionisation des différents acides aminés ionisables (Asp, Glu, Arg, Lys et His). Cela est réalisé en utilisant les définitions du champ de force AMBER99SBildn (aujourd'hui intégrées dans la version SB14 du champ de force d'AMBER pour les protéines). Le choix du champ de force pour le paramétrage est un autre point de décision important puisque les outils informatiques utilisés (AmberTools) donnent accès à différentes versions des champs de forces. Dans le cas de hPRDX5, il est critique de bien prendre en compte l'état d'ionisation en thiolate à pH neutre de la cystéine catalytique Cys47 (Hall et al. et références citées).¹⁰⁶ AMBER permet de définir un thiolate en déclarant un acide aminé de nomenclature « CYM » au lieu de « CYS » dans le fichier pdb de la protéine.

- Création et préparation du ligand

Une fois le fichier de structure de la protéine préparé, il faut préparer les fichiers de structure des ligands. Pour cela la construction du ligand peut-être effectuée à partir de données cristallographiques, ou de novo, par exemple avec le logiciel Avogadro.¹⁶⁷ On peut ainsi construire de quelconques ligands et en effectuer une minimisation en énergie initiale. A partir des coordonnées créées, la paramétrisation du ligand est nécessaire pour l'obtention des paramètres compatibles avec ceux du champ de force d'AMBER. Une approche classique est l'usage du champ de force généralisé d'AMBER appelé GAFF (*Generalized AMBER Force Field*). Pour cela on utilise la suite d'outils Antechamber¹⁶⁸ qui va permettre d'identifier les atomes et liaisons du ligand, d'ajouter les hydrogènes et de créer le fichier de charges électriques partielles associé (.prepi). L'utilitaire « parmchk » permet ensuite de vérifier si tous les paramètres requis sont disponibles, et va produire un fichier avec les éventuels paramètres de champ de force manquants associés à GAFF. Le ligand (éventuellement dupliqué en plusieurs entités si besoin de définir une concentration plus élevée) est ensuite disposé par rapport à la protéine pour définir le système à solvater.

- Solvation et addition d'ions dans le système

La solvation du système nécessite le choix du modèle de description de l'eau (TIP3P ici) et des ions monovalents (Na⁺ et Cl⁻).⁵⁵ Dans Antechamber, l'étape de solvation et d'ionisation du système est effectuée par le programme « tleap ». On contrôle aussi la définition de la boîte d'eau périodique, les sels et leurs concentrations pour obtenir la neutralité électrique globale. Une vérification de la charge globale et du nombre d'ions Na⁺ et Cl⁻ à ajouter pour obtenir un système

électriquement neutre est réalisée par des scripts du logiciel VMD. Afin de mimer une solution de hPRDX5 à 150 mM NaCl qui sera proche des conditions expérimentales en RMN et en cinétique enzymatique, on ajoute typiquement 55000 molécules d'eau, 48 ions Na^+ et 46 ions Cl^- (pour neutraliser la charge globale de -2 de hPRDX5) dans une cellule d'eau de dimension d'environ 83 x 83 x 95 Å avec la protéine centrée (Figure 41).

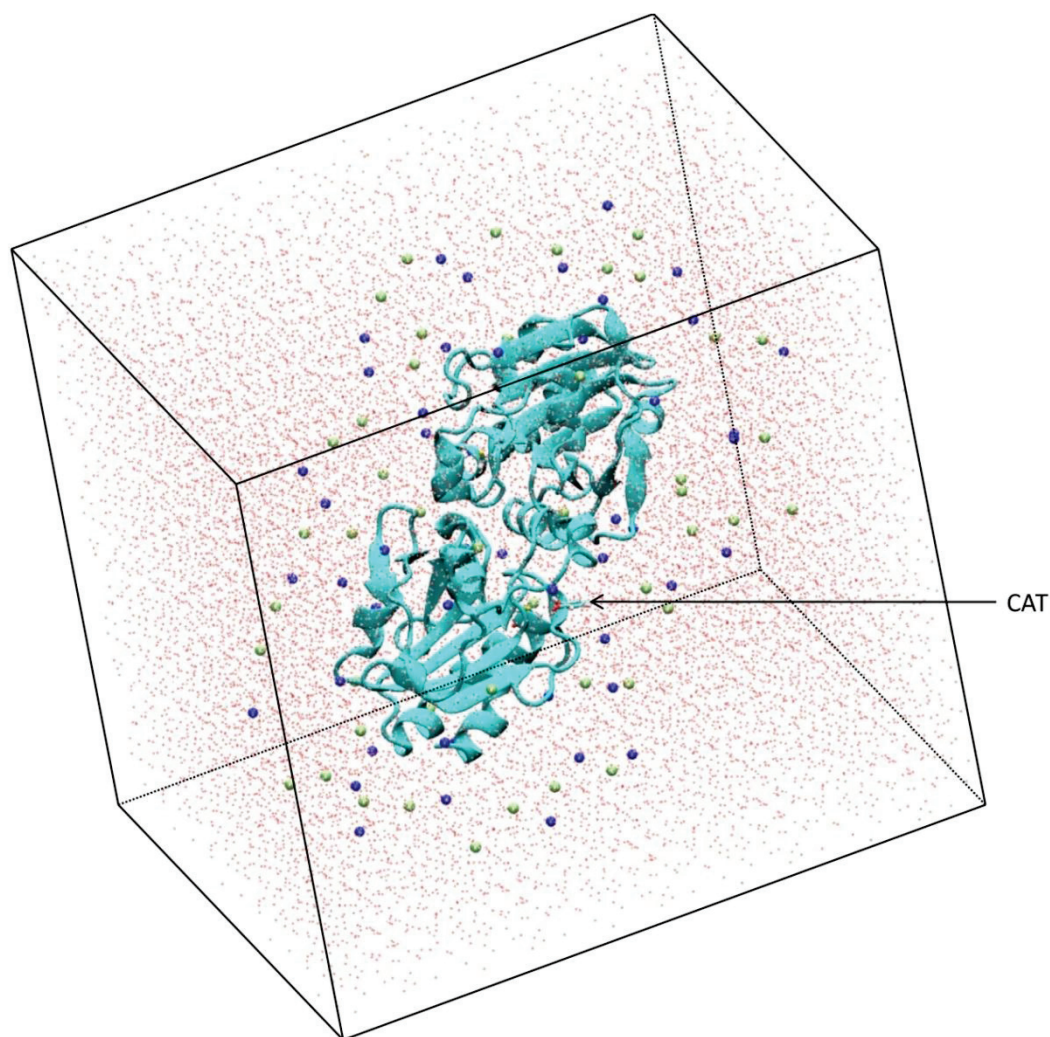


Figure 41. Visualisation du système hPRDX5-Catechol utilisé en MD et FM, après les étapes de solvatation et d'ajout des sels.

Structure de hPRDX5 visualisée avec VMD (cyan), les molécules d'eau en points fins, les ions Na^+ et Cl^- sous forme de sphères de van der Waals (violet et vert respectivement) et le ligand CAT et la cystéine catalytique Cys47 sous forme de bâtons.

I. 1. 1 Processus d'équilibration

Avant d'initier une production de MD, il est indispensable d'effectuer une étape d'équilibration du système. En effet, comme les molécules d'eau sont disposées de façon géométrique régulière et les ions de façon équirépartie, il est essentiel d'amener le système à un point sans conflit énergétique infini. Le protocole choisi est le même que celui utilisé pour l'équilibration de systèmes membranaires plus complexes. Coûteux en calcul (1 à 2 h sur 1 GPU), il a été choisi en vue de laisser le temps à l'eau de diffuser dans la protéine lorsque le modèle initial ne comportait pas d'eau, et fût ensuite choisi comme standard opératoire. Le principe général consiste à contraindre spatialement les atomes lourds de la protéine et de laisser l'eau, les sels, et les hydrogènes de la protéine s'adapter progressivement, dans un processus de recuit simulé où les contraintes spatiales sont relâchées par étapes avant d'être réimposées (voir partie I.1 du chapitre II).

Au cours de l'équilibration, on suit l'évolution de différents paramètres: la pression du système (bar), sa pression moyenne au cours du temps (bar), sa température (K), et la force de contrainte appliquée aux atomes lourds ($\text{kcal.mol}^{-1}.\text{\AA}^{-2}$) (Figure 42).

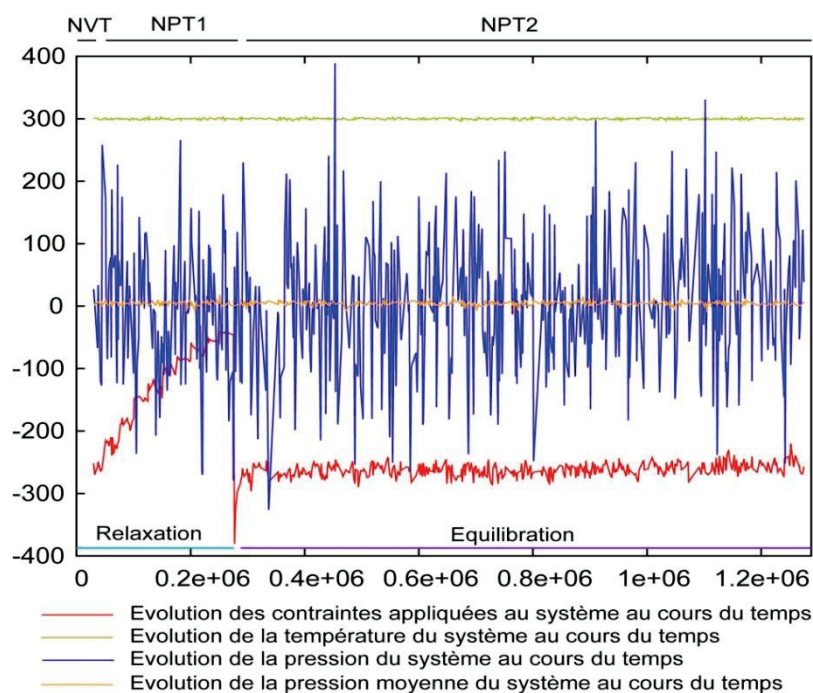


Figure 42. Graphique représentant l'évolution des différents paramètres indiqués du système en fonction du temps, durant l'étape d'équilibration d'un système hPRDX5 en solvant explicite (TIP3P) étudié par MD. Le système est laissé à relaxer durant les périodes NVT et NPT1, et est équilibré durant la période NPT2.

Il est important de vérifier que la pression moyenne du système avoisine 1 bar et que sa température soit stable au cours du temps (fixée ici à 300K). A la fin du processus, le système est en principe prêt pour initier un calcul de MD.

I. 1. 2 Lancement d'une dynamique non biaisée

La seconde phase préalable d'un projet de MD est d'étudier l'évolution du système au cours d'une trajectoire de dynamique prolongée sans contraintes (typiquement d'une durée de 100 à 500 ns lorsque l'on dispose de GPU). Cette étape peut être effectuée avec ou sans ligand. En présence de ligand elle peut aussi représenter une première analyse de l'interaction PL si on y observe la formation d'un complexe ou sa dissociation, et permettre la détermination des variables collectives à utiliser pour suivre le phénomène en métadynamique. Des simulations de type NVT sans contraintes sont typiquement effectuées en utilisant les protocoles définis en chapitre II (I.2 p.78).

Un de nos protocoles préalable est celui où la protéine est en présence de 6 ligands distants autour de la protéine pour une concentration totale de 6 mM environ dans la boîte d'eau modélisée. Certains de ces calculs ont même été effectués en présence du DTT oxydé dans le site récepteur, afin de voir si les ligands pouvaient entrer en compétition avec. Des calculs de 500 ns ont été effectués pour l'analyse des événements de liaison et de dissociation des ligands avec la protéine (voir en partie I.2.1 de ce chapitre).

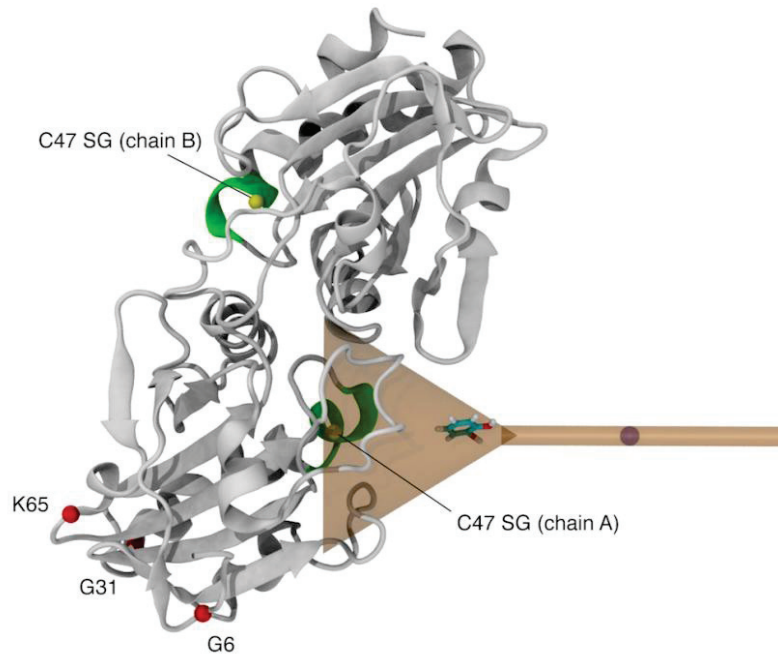
I. 1. 3 Mise en place d'un calcul de FM

Une fois que la simulation NVT a montré une possible liaison du ligand au site récepteur de l'enzyme, la mise en place de la FM est possible. Pour ceci, il faut dans un premier temps calculer le potentiel externe en forme d'entonnoir utilisé, et définir les variables collectives (CV) à utiliser dans un protocole utilisant le plugin PLUMED (version 1.3).⁶⁴

- Paramétrage du potentiel en entonnoir

Le potentiel en forme d'entonnoir est calculé selon un référentiel fixe dans l'espace qui sera appelé dans le protocole PLUMED. Le calcul de ce potentiel est réalisé par un *shell script* selon V. Limongelli et coll..⁷² Lors des premières simulations, l'axe Z du potentiel avait pour origine l'atome de soufre de la cystéine catalytique SG Cys47, et était orienté en passant par l'un des atomes lourds du ligand défini pour chaque simulation (Figure 43). Cette orientation dépendait donc de la position initiale du ligand à chaque simulation.

A)



B)

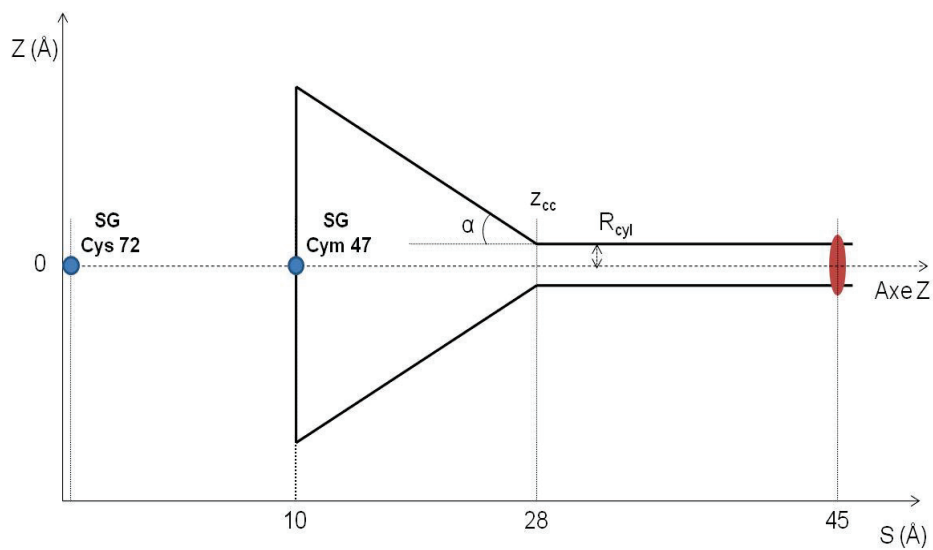


Figure 43. Paramétrage du potentiel de contrainte spatiale du ligand utilisé en FM.

Visualisation du potentiel appliqué au niveau du site actif du dimère de hPRDX5 en présence d'un ligand catechol. Le thiolate de la cystéine catalytique SG Cys47 est représenté par une sphère jaune sur les deux chaînes de l'homodimère. Les atomes C α des résidus Gly6, Gly31 et Lys65 sont représentés par des sphères rouges, et sont utilisés comme contraintes de position pour inhiber la diffusion globale de la protéine. La partie cylindrique du potentiel est orientée selon l'axe Z avec un rayon de 1 Å, en passant par SG Cys47 et l'un des atomes de carbone du cycle aromatique du ligand. Le potentiel possède un mur initialement situé à 35 Å sur l'axe Z. B) Schéma montrant la normalisation de l'orientation du potentiel en entonnoir utilisé avec l'axe Z passant par SG Cys72 et SG Cys47.

Afin d'avoir toujours la même orientation, nous avons choisi de définir comme origine de l'axe Z l'atome de soufre de la cystéine 72 (SG Cys72) présent dans les feuillets β de la protéine, et d'orienter cet axe en passant par l'atome de SG Cys47, situé à environ 10 Å de l'origine. Les paramètres utilisés, représentés en Figure 42, sont : le rayon du potentiel cylindrique $R_{cyl} = 1\text{Å}$; la distance sur l'axe Z à laquelle le potentiel change de forme $Z_{cc} = 28\text{Å}$; l'angle formé par le potentiel en forme de cône $\alpha = 0.55\text{ rad}$; les paramètres de progression $\text{min}_S = 8\text{Å}$ et $\text{max}_S = 45\text{Å}$; et les paramètres de distance $\text{min}_Z = 0$ et $\text{max}_Z = R_{cyl} + \tan \alpha * (Z_{cc} - \text{min}_S)$.

- Définition des variables collectives (CV)

Comme présenté en partie d'introduction, les CV doivent permettre de pouvoir décrire le système simulé, et de distinguer les différentes conformations liées du ligand. De manière générale, 2 ou 3 CV sont suivies au cours d'un calcul de métadynamique grâce au plugin PLUMED.⁶⁴ Le biais a été appliqué sur CV1 (la distance entre le ligand et la protéine à l'intérieur du potentiel suivant l'axe Z), tandis que CV2 (la distance du ligand par rapport à ce même axe Z) et CV3 (l'angle dièdre SG-O_A-C₁-C₂ formé par l'atome SG de Cys47, un des deux atomes d'oxygène du ligand, et les deux atomes de carbone voisins) ont été suivies au cours du temps de simulation.

Cependant, nous nous sommes rendu compte que le suivi des torsions dièdres SG-O_A-C₁-C₂ ne permettait pas de discriminer les différentes conformations de liaison du complexe PL. Nous avons donc modifié CV3 en choisissant de suivre plutôt la valeur de l'angle SG-O_A-O_B formé par l'atome SG de Cys47 et les deux atomes d'oxygène du ligand (O_A et O_B) (Figure 44). Cette nouvelle définition permet de pouvoir corréler les conformations de liaison en fonction du nombre de liaisons hydrogènes formées entre le ligand et Cys47 puisqu'elle permet d'avoir des informations sur les deux atomes d'oxygène.

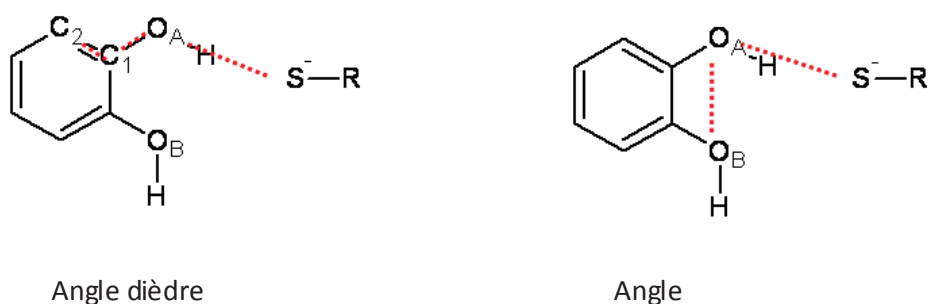


Figure 44. Définition de la variable collective CV3 suivie en FM.

CV3 est définie sous forme de torsion SG-O_A-C₁-C₂, puis ensuite sous forme d'angle SG-O_A-O_B, pour la caractérisation du système protéine ligand. R = Cys47

I. 2. Détermination de l'affinité protéine-ligand par FM

L'avantage considérable de la FM est à la fois l'exploration des événements de liaison et de dissociation du complexe PL, mais aussi l'obtention de l'affinité convergée grâce à la répétition de ces événements. Nous allons décrire ci-dessous les différentes étapes qui mènent à ce résultat.

I. 2. 1 Dynamique non biaisée NVT

Le catéchol, découvert initialement comme *hit* se liant au site actif de hPRDX5 par criblage RMN,⁹¹ nous a permis de tester la simulation de dynamique, et de vérifier qu'effectivement un ligand (au moins) allait se lier spontanément (sans biais) à l'une des deux cystéines catalytiques des sous-unités A et B de l'homodimère de hPRDX5. Une simulation de 500 ns de dynamique NVT sans contraintes a été complétée. Six catéchols et une molécule de DTT ont été mis en compétition, nommés ici BRG, CRG, DRG, ERG, FRG et GRG (Figure 45).

La Figure 44 montre que la molécule de DTT se détache de la protéine après environ 100 ns de simulation (flèche violette). En tout, 3 événements de liaison (de courte durée) seulement sont enregistrés sur l'ensemble des 6 ligands (flèches bleues), en prenant en compte les 2 sites actifs de l'homodimère de hPRDX5. Durant cette période de liaison du ligand à hPRDX5, le ligand se lie principalement par une seule liaison hydrogène, formant un angle $SG-O_A-O_B$ de 180° entre SG de Cys47 et les deux atomes d'oxygène du ligand concerné. Cependant, il a aussi été observé que le ligand peut faire une deuxième liaison hydrogène grâce au second groupe hydroxyle du catéchol, et ainsi se lier à Cys47 avec un angle $SG-O_A-O_B$ de 60° cette fois comme visible en Figure 45, à 30 ns sur le graphique représentant le ligand BRG. Ce changement de conformation peut se faire *in situ* (sans avoir à repasser par la forme libre) car le ligand revient ensuite à une conformation de liaison avec une seule liaison hydrogène.

De la même manière, en tout début de simulation le ligand GRG se lie à Cys47 (chaîne B) d'abord par une seule liaison hydrogène, puis deux, et enfin revient à la conformation liée par simple liaison H avant de se détacher.

Les autres ligands diffusent librement dans le solvant, et on peut distinguer des interactions de surface avec hPRDX5 comme par exemple les ligands BRG et FRG durant une période non négligeable.

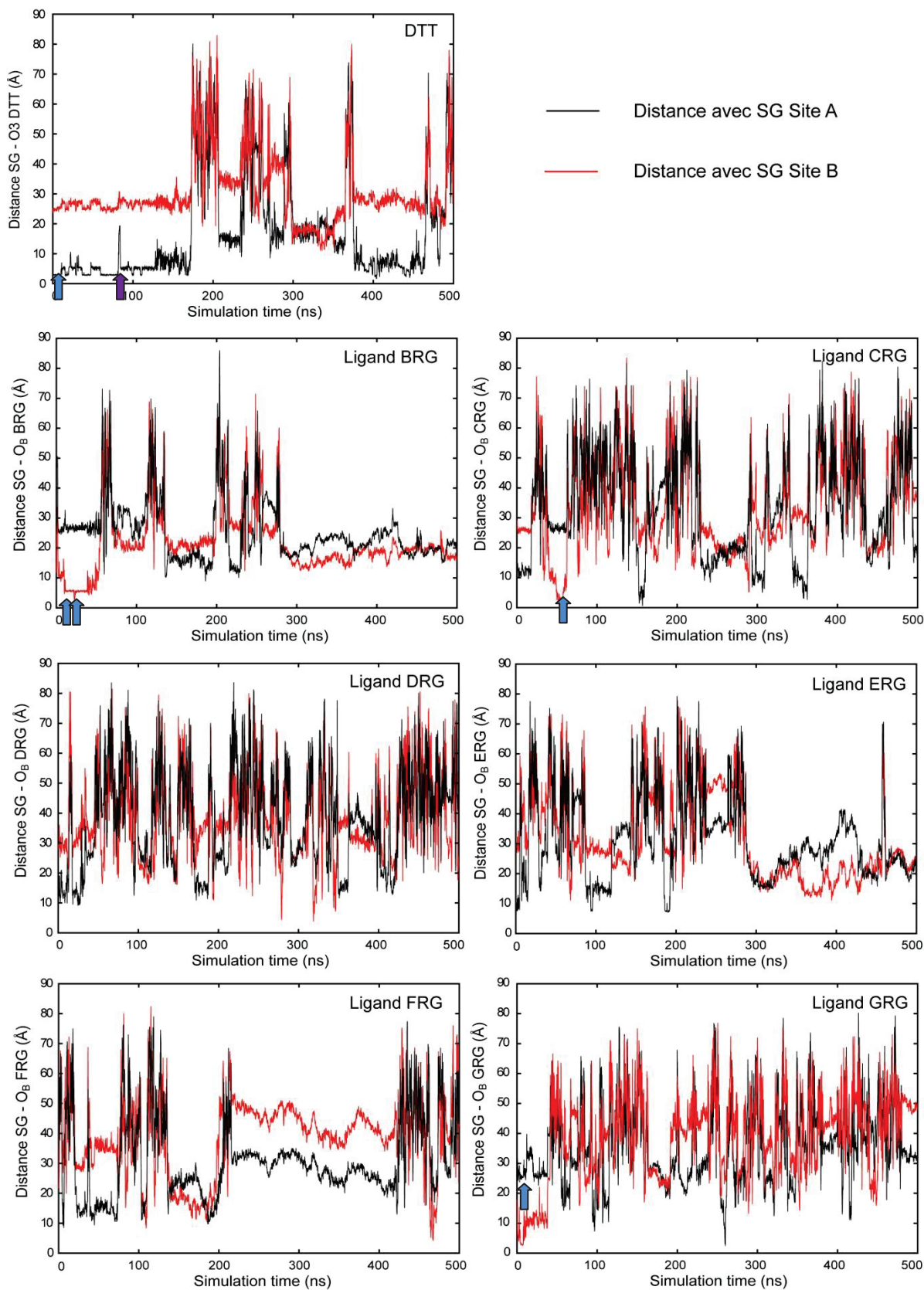


Figure 45. Dynamique NVT de hPRDX5 complexée au DTT oxydé en présence de 6 catéchols.

Tracé des distances entre l'atome d'oxygène O_{β} de chacun des 6 CAT (nommés BRG, CRG, DRG, ERG, FRG et GRG) et l'atome de soufre SG des 2 cystéines catalytiques des sites actifs des chaînes A et B (Å). Les flèches bleues indiquent un événement de liaison et la flèche violette indique la dissociation du DTT.

Par cette simulation, nous avons pu réaliser que le CAT allait effectivement se lier spontanément à la cystéine, selon deux modes de liaison représentés en Figure 46. Le premier mode de liaison est comparable à celui observé en cristallographie aux rayons X dans les structures de hPRDX5 en présence de ligand tel que le DTT oxydé dans le site actif.¹⁰⁶ La seconde conformation avec une double liaison hydrogène formant un angle $SG-O_A-O_B$ de 60° n'a par contre jamais été observée de manière expérimentale. Selon une étude théorique de l'oxydation des thiols par H_2O_2 effectuée en 2005, cette conformation pourrait mimer le complexe enzyme-substrat stabilisé grâce à un changement conformationnel permettant la liaison des deux groupements hydroxyles, avant de réagir avec le groupement thiolate.¹⁸⁵

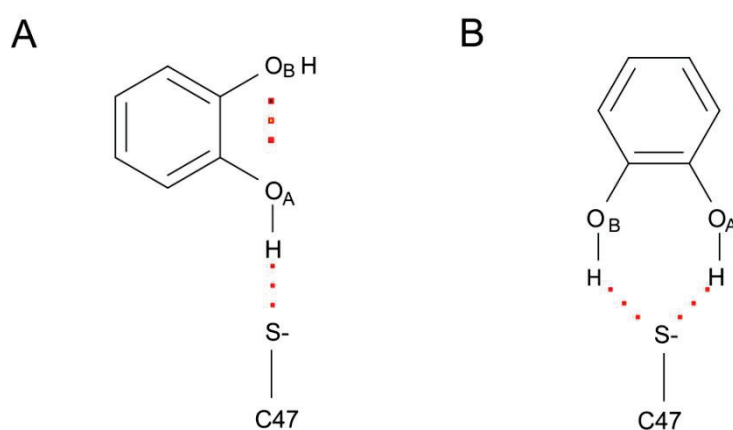


Figure 46. Les deux modes de liaison du catéchol observés en dynamique NVT avec hPRDX5.

A) Angle $SG-O_A-O_B$ de 180° , qui reflète une conformation alignée à 1 liaison H avec SG Cys47 B) Angle $SG-O_A-O_B$ de 60° qui reflète une conformation doublement liée à SG Cys47 par liaisons H.

Dans le cas du ligand 4MEC, le DTT se dissocie du site catalytique après seulement 10 ns de simulation, et un des 6 ligands vient rapidement prendre sa place dans la cavité laissée libre. Il se positionne en liaison H avec le thiolate de la cystéine de la chaîne A (à environ 40 ns). Ce ligand va rester en liaison durant 260 ns, se dissocier rapidement avant de se lier de nouveau mais de manière brève pour finalement se dissocier de manière définitive et retourner diffuser dans le solvant.

Ces évènements de liaison et de dissociation ont un caractère aléatoire. Au niveau des 5 autres ligands simulés, plusieurs évènements de liaison sont visibles, dans chacun des deux sites (4 évènements dans le site actif de la chaîne B, et un autre évènement dans la chaîne A). En tout, 6 évènements de liaison sont comptabilisés, au cours desquels on distingue deux types de conformations liées (Figure 47). La première, majoritaire en termes de présence comme dans le cas de CAT, est la conformation où le ligand est lié par l'un des hydroxyles à la cystéine catalytique par le

biais d'une liaison hydrogène (angle $SG-O_A-O_B$ de 180°). La deuxième conformation observée lie le ligand à la cystéine par le biais de deux liaisons hydrogènes formées par les hydroxyles, formant alors un angle $SG-O_A-O_B$ de 60° environ comme observé précédemment dans le cas du CAT. Nous avons de plus observé que ces deux conformations pouvaient s'inter-convertir sans avoir à se dissocier.

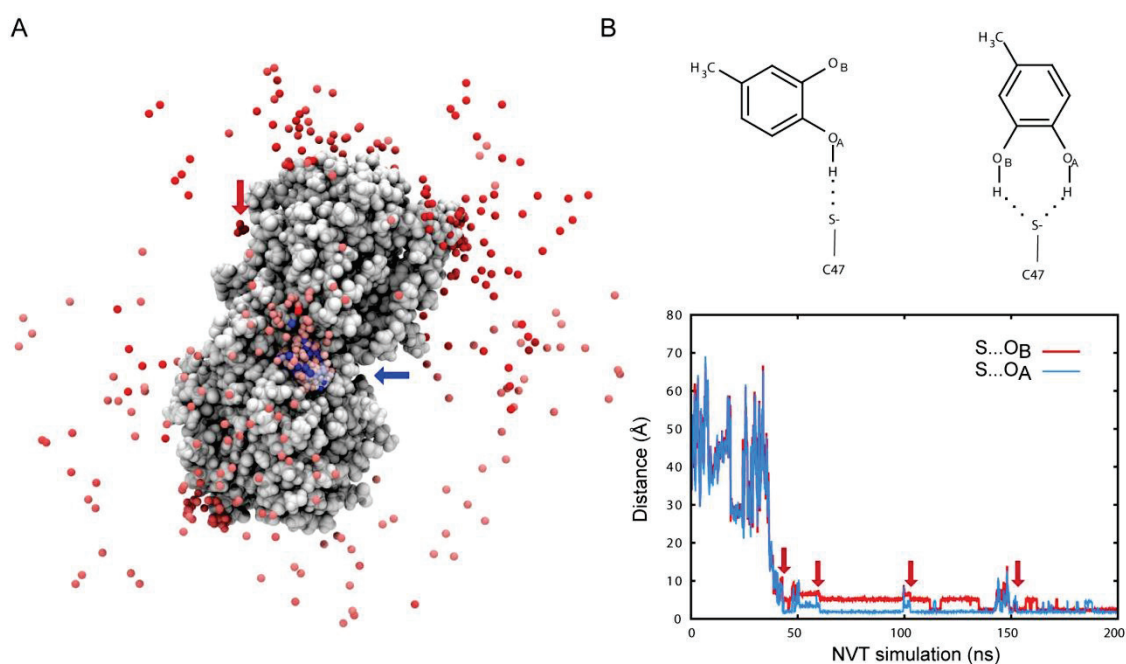


Figure 47. Analyse de la trajectoire NVT de hPRDX5 complexé au DTT oxydé en présence de 6 ligands 4-méthylcatéchol au cours de la dynamique NVT.

A) Aperçu de la trajectoire d'un des 6 ligands 4-méthylcatéchol (4MEC) au cours de la dynamique NVT (de rouge vers bleu). B) Distances (Å) entre l'atome SG de Cys47 (chaîne A) et les atomes d'oxygène O_A et O_B d'une molécule de 4MEC. Evènements de liaison signalés par une flèche rouge.

I. 2. 2 Funnel-métadynamique (FM)

Les interactions de CAT et de 4MEC avec hPRDX5 ont ensuite été étudiées par des simulations de FM de 500 ns. Dans ce protocole, un seul CAT et un seul 4MEC sont utilisés, et disposés au sein du potentiel externe en entonnoir. Pour chacun des ligands, plusieurs évènements de liaison ont été observés (entre 8 et 10).

La statistique des différentes variables collectives suivies au cours de la simulation a été collectée, et peut être visualisée sur la Figure 48.

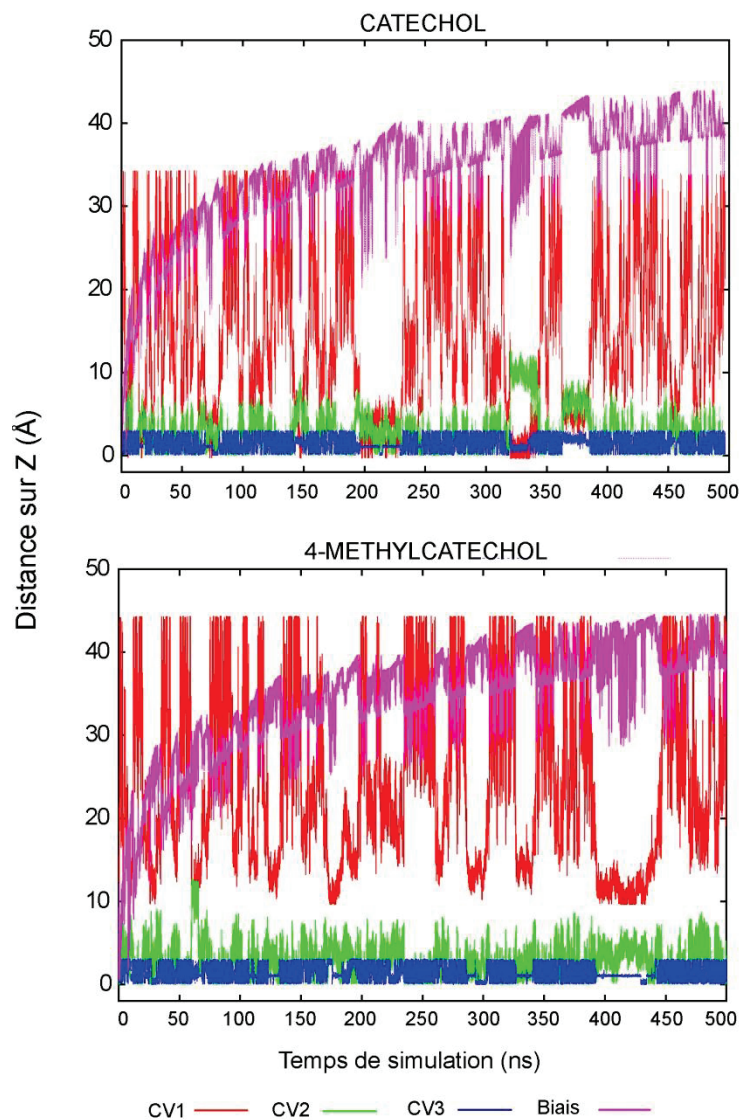


Figure 48. Evolution des différentes variables collectives au cours du temps de simulation FM des systèmes CAT-hPRDX5 (haut) et 4MEC-hPRDX5 (bas).

CV1 (distance sur l'axe Z) représentée en rouge, CV2 (distance par rapport à l'axe Z) en vert, CV3 (angle SG-O_A-O_B) en bleu, et le dépôt du potentiel de HILLS (biais) en rose. La distance du centre de masse du ligand sur l'axe Z dans le cas de CAT est compris entre 0 et 35 Å, tandis qu'elle varie entre 10 et 45 Å dans le cas du 4MEC car la définition de l'origine de l'axe Z du potentiel en entonnoir a été modifiée comme expliqué au début de ce chapitre.

CV1 représente la distance du centre de masse du ligand sur l'axe Z, CV2 la distance par rapport à ce même axe et CV3 représente la valeur de l'angle SG-O_A-O_B formé. Le biais appliqué au système sous forme de potentiels de HILLS est aussi visible sur ce graphique. On voit que la valeur de l'angle formé entre SG Cys47 et les deux oxygènes du ligand O_A et O_B, reste stable durant les périodes de liaison (Figure 48). Cependant, il y a de brusques modifications de cet angle durant les périodes de liaison, représentant le changement de conformation du ligand en liaison avec la cystéine, observé précédemment en dynamique NVT. Cet angle de 180° dans le cas de la liaison hydrogène d'un seul groupe hydroxyle à SG Cys47, passera instantanément à 60° si le deuxième groupe hydroxyle vient se

lier à son tour. Une transition vers le second mode de liaison impliquant les deux hydroxyles se produit dans le cas du CAT à 220 ns de simulation et dans le cas du 4MEC à 440 ns de simulation, avec retour rapide à la conformation liée par simple liaison H.

Il est possible de représenter le potentiel de force moyen (PMF) $W(Z)$ du système en utilisant la valeur du potentiel de biais en fonction d'une ou plusieurs CVs, au cours du temps de simulation, à l'aide d'un algorithme de repondération (V. Limongelli).⁷² Il est donc possible de visualiser par exemple le PMF du système en fonction de la distance du centre de masse du ligand sur l'axe Z (CV1) (Figure 49). Le minimum énergétique correspond à la formation du complexe PL, et le plateau avec l'énergie maximum correspond à l'état libre. Il est aussi possible de reconstruire des cartes du PMF en le représentant en fonction de 2 CV (FES pour *Free Energy Surface*), ce qui permet une description plus complète de l'interaction. La Figure 48 représente les FES des systèmes CAT-hPRDX5 et 4MEC-hPRDX5 en fonction de CV1 et CV2. On voit que la totalité (ou quasi totalité pour 4MEC) de l'espace conformationnel disponible est échantillonné à l'intérieur du potentiel en entonnoir.

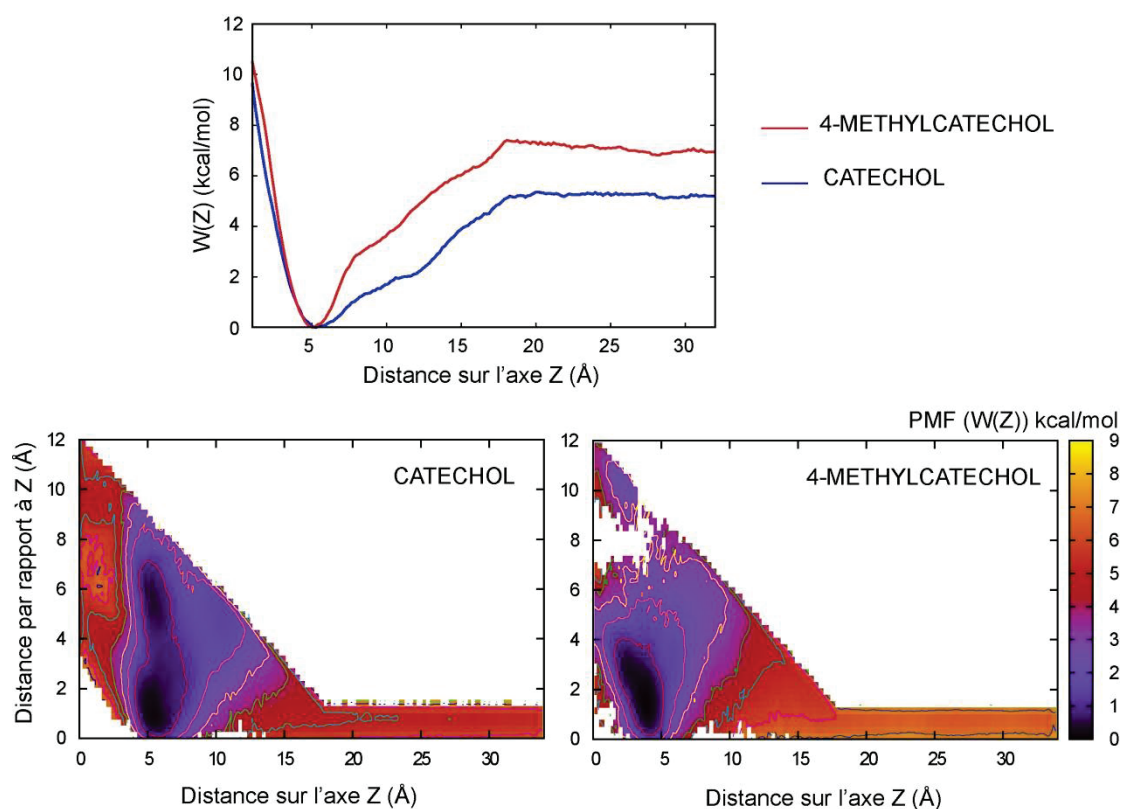


Figure 49. Courbes $W(Z)=f(CV1)$ et surfaces $W(Z)=f(CV1,CV2)$ des complexes hPRDX5-CAT et hPRDX5-4MEC étudiés par FM.

Valeur du potentiel $W(Z)$ en kcal/mol en fonction de CV1 (distance sur l'axe Z) ou de CV1 et CV2 (distance par rapport à l'axe Z). $W(Z)$ représenté sur les surfaces avec des courbes de niveau d'intervalle 1kcal, et une surface de couleur allant du noir (énergie la plus basse) au jaune (énergie la plus haute).

Si l'on considère la courbe de PMF en fonction de CV1, on remarque que la différence d'énergie entre l'état libre dans le solvant et l'état lié (minimum énergétique) est plus importante d'1 kcal dans le cas du 4MEC (7 kcal.mol⁻¹) que pour le CAT (6 kcal.mol⁻¹). Les FES étant uniquement construites en fonction de CV1 et CV2, on remarque qu'il est impossible de différencier dans ce cas les 2 conformations de liaison observées en dynamique NVT, qui ne diffèrent que par la valeur de l'angle SG-O_A-O_B. Pour les deux ligands un seul bassin d'interaction comprend donc les 2 conformations de liaison. C'est pour cette raison que nous avons aussi suivi l'angle SG-O_A-O_B au cours de la simulation (voir plus loin Figure 52).

La convergence de l'énergie a été suivie par la valeur de dépôt du biais au cours de la simulation (Figure 50), et confirmée par le tracé de l'énergie ΔG°_b en fonction du temps, calculée à partir de l'équation 26 présentée en chapitre d'introduction selon V. Limongelli (Figure 51).

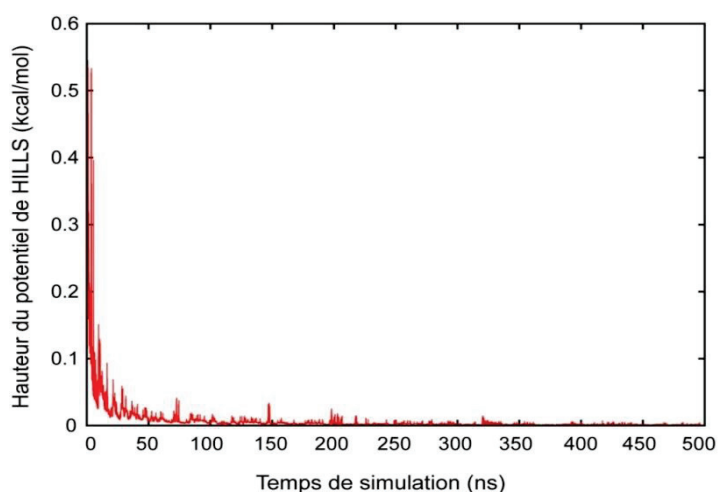


Figure 50. Dépôt du potentiel de biais HILLS au cours du temps de simulation .

Graphique obtenu dans le cas du ligand CAT. Dépôt du potentiel de biais durant 500 ns de simulation FM : dépôt nul, signifiant que l'énergie est convergée

La différence d'énergie du système qui oscille entre un état libre et lié a été moyennée (moyenne pondérée) en fonction du temps, et calculée sur les 100 dernières nanosecondes de simulation, où on est à la convergence. Le résultat est exprimé avec la fluctuation associée de cette énergie en fin de simulation, sur cette même période de temps et le K_D calculé selon l'équation 5 (p.8).

CAT	$\Delta G^{\circ}_b = - 3.1 \pm 0.2 \text{ kcal.mol}^{-1}$	$K_D = 6.1 \pm 1.8 \text{ mM}$
4MEC	$\Delta G^{\circ}_b = - 4.2 \pm 0.3 \text{ kcal.mol}^{-1}$	$K_D = 0.9 \pm 0.4 \text{ mM}$

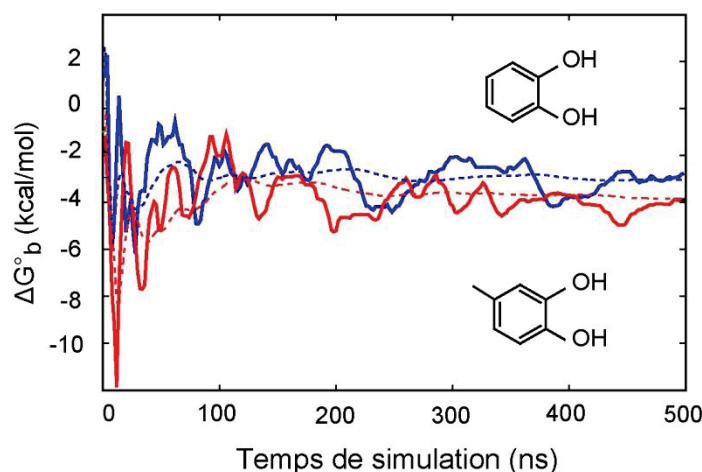


Figure 51. Evolution de l'énergie libre de liaison (kcal.mol⁻¹) au cours des 500 ns de simulation FM pour les systèmes CAT-hPRDX5 et 4MEC-hPRDX5.

L'énergie libre de liaison (ΔG°_b) est représentée en trait plein, et l'énergie libre moyenne pondérée au cours du temps en pointillé. Catéchol en bleu, et 4-méthylcatéchol en rouge.

I. 2. 3 Vue détaillée de l'interaction protéine-ligand par FM

Nous avons vu que la dynamique NVT indique que lorsque le ligand occupe le site actif, deux conformations de liaison sont en échange (angle $SG-O_A-O_B$ de 60° et de 180° entre SG Cys47 et les atomes d'oxygène du ligand) (voir Figure 46). En cartographiant l'énergie potentielle des systèmes hPRDX5-CAT et hPRDX5-4MEC en fonction de CV1 et CV2 (Figure 49) on ne peut discriminer ces conformations. Mais si on cartographie l'énergie libre en fonction de l'angle $SG-O_A-O_B$, il est alors possible à travers la surface d'énergie libre (FES) obtenue de différencier les deux bassins d'interaction A et B, et d'évaluer la stabilité de leur état lié par rapport à leur différence en énergie avec l'état libre (Figure 52). Le conformère A observé est comparable aux conformations décrites par cristallographie dans le cas d'autres ligands comme le benzoate, ou bien le DTT oxydé.¹⁰⁶

Dans le cas du CAT, chaque bassin va correspondre à une conformation du fait de la symétrie présente dans cette molécule. Pour un angle $SG-O_A-O_B$ de 0° (ou 180°), les 3 atomes seront alignés et l'interaction se fera par simple liaison hydrogène (bassin A) entre l'un des 2 oxygènes du ligand, et SG Cys47 (Figure 52). Le bassin visible autour de 120° correspond lui aussi à une liaison hydrogène simple, mais avec une orientation différente du noyau aromatique. Pour un angle $SG-O_A-O_B$ de 60° , les deux groupes hydroxyles sont en liaison H avec le thiolate, comme vu précédemment en dynamique NVT (Figure 46). L'énergie des 2 bassins est équivalente. Dans le cas du 4MEC par contre, il n'est pas possible de relier les bassins A et B à une seule conformation car la molécule ne possède pas d'axe de symétrie.¹⁸³ Pour chaque angle donné, 2 conformations sont possibles selon l'orientation du groupement méthyle (Figure 52).

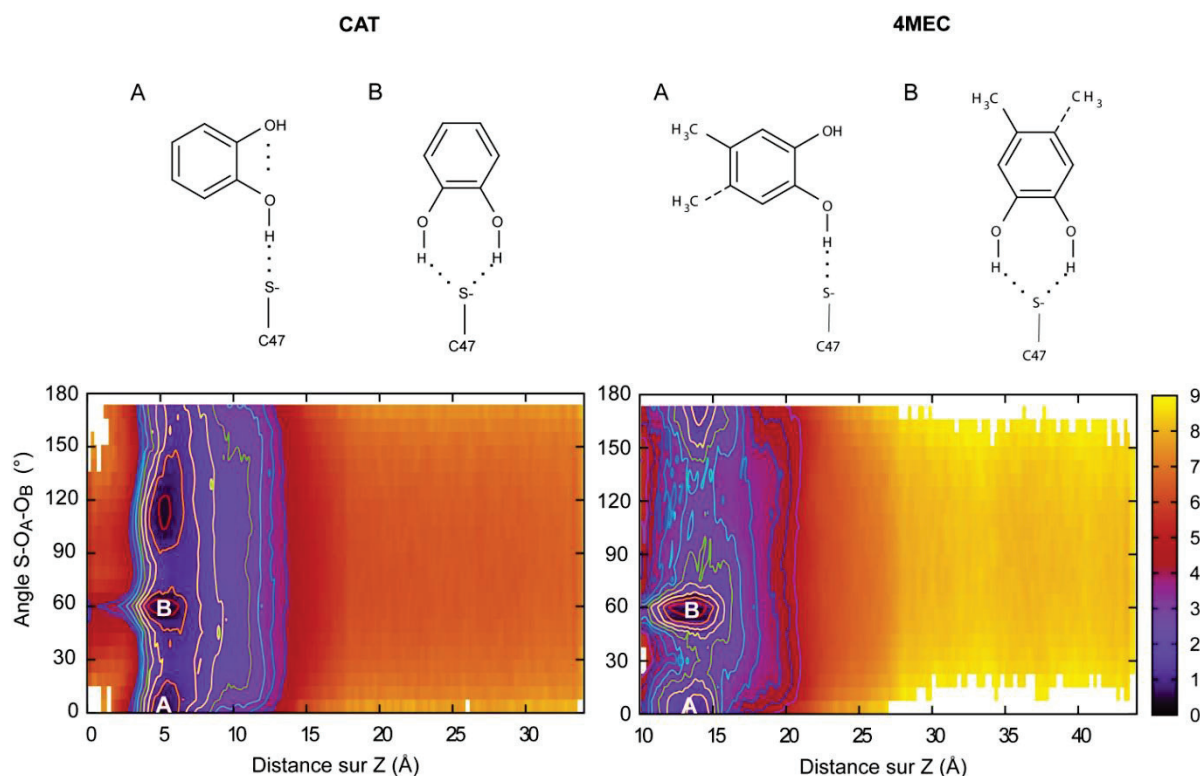


Figure 52. Etude détaillée de l'interaction de CAT et 4MEC avec hPRDX5 par FM.

Cartographie de l'énergie libre du complexe PL en fonction des variables collectives 1 (distance sur l'axe z) et 3 (angle SG-O_A-O_B). Contours iso-énergétiques toutes les 0.5 kcal du minimum jusque 5 kcal.mol⁻¹.

Ces 4 conformations possibles sont celles que l'on retrouve en effectuant le *clustering* géométrique de l'ensemble de la trajectoire, à l'aide du logiciel de visualisation de dynamique moléculaire VMD avec une valeur seuil de 2 Å (Figure 53). Les deux conformations majoritaires sont les conformations B et C avec le groupement méthyle orienté vers l'intérieur de la cavité du site actif, les états A et D appartenant respectivement aux bassins A et B, étant des conformations que l'on retrouve mais de manière minoritaire. Ces modes de liaison sont en accord avec les résultats publiés en 2014 dans lesquels les ligands CAT et 4MEC sont étudiés par cristallographie en interaction avec hPRDX5.¹⁸²

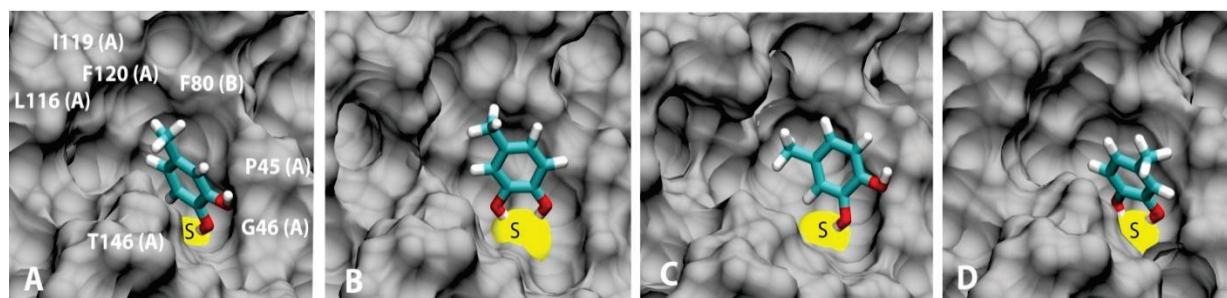


Figure 53. Clustering géométrique des complexes hPRDX5-4MEC obtenus par FM.

Visualisation des 4 conformations identifiées avec un cutoff de distance de 2Å, correspondant aux 2 bassins d'interaction A et B visibles en Figure 47. Cys47 représentée en jaune.

Dans le cristal, c'est la conformation où le ligand est en simple liaison H avec SG de Cys47 qui est observée, correspondant au bassin A des FES. De plus dans le cas du système hPRDX5-4MEC, le groupement méthyle est orienté vers l'hélice hydrophobe Leu116 - Phe120, ce que l'on retrouve dans 3 des 4 clusters identifiés par FM (Figure 53).

I. 3. Corrélation de l'affinité calculée par FM et déterminée par RMN

Nous avons décidé d'étudier par RMN l'affinité des systèmes hPRDX5-catéchols afin de valider expérimentalement les résultats d'affinité obtenus par dynamique. Nous avons vu précédemment que la RMN est une méthode très adaptée pour la détermination de l'affinité.

I. 3. 1 Détermination de l'affinité protéine-ligand par RMN

Pour déterminer expérimentalement l'affinité de CAT et 4MEC pour la forme réduite de hPRDX5 et corrélérer avec les résultats d'affinité obtenus par FM, nous avons utilisé des expériences ¹⁵N-¹H HSQC. Nous avons mesuré les perturbations de déplacement chimique (CSP) induites par CAT et 4MEC en fonction de leur concentration, sur les acides aminés de hPRDX5. Le DTT qui est souvent utilisé comme agent réducteur de ponts disulfures de protéines n'a pas été utilisé ici car sa forme oxydée mentionnée précédemment comme ligand cristallographique est à priori susceptible d'induire une compétition avec les ligands.¹⁰⁶ Nous avons donc préféré utiliser un réducteur de disulfures non apparenté au DTT : le TCEP, qui n'induit aucune variation de déplacement chimique détectable lorsqu'il est utilisé seul avec hPRDX5, indice apparent de non-affinité du réducteur ou de son produit d'oxydation pour hPRDX5.

Lorsque l'on expose la protéine ¹⁵N-hPRDX5 à des concentrations croissantes de ligand, on observe des CSPs sélectives de certains signaux RMN. L'attribution de ces signaux étant par ailleurs disponible au laboratoire (I. Krimm, non publié), nous pouvons non seulement déterminer l'affinité apparente des ligands pour hPRDX5, mais aussi avoir une connaissance des acides aminés affectés. Pour chaque concentration en ligand, un spectre HSQC a été enregistré et la variation de déplacement chimique calculée par rapport au spectre de la protéine seule. Les acides aminés présentant les CSPs les plus significatives ont été retenus pour l'analyse. Les CSPs sont analysées en fonction du rapport des concentrations en ligand par rapport à la concentration en protéine ($[L]/[P]$). La valeur du K_D est déterminée par régression non linéaire sous MatLab (chapitre II p.95) (O. Walker).¹⁷⁸ Les ligands étant apportés en utilisant une solution concentrée dans le DMSO, nous avons vérifié expérimentalement que la présence de DMSO n'avait pas d'influence significative sur le déplacement chimique des signaux de la protéine jusqu'à un seuil de 10% v/v (Figure 54). La protéine n'a subi

aucune dénaturation ou modification conformationnelle après ajout de 10% de DMSO. Pour l'ensemble de l'étude, la quantité de DMSO maximale introduite ne dépasse pas 8%.

Dans le cas du CAT, la titration a été effectuée jusqu'à atteindre un ratio $[L]/[P]$ de 25, correspondant à une concentration maximale en catéchol de 10 mM pour $[P]_0 = 400 \mu\text{M}$. Les spectres ont été superposés afin de visualiser les perturbations de déplacement chimique induites par le ligand sur les signaux de hPRDX5, qui ont ensuite été mesurées graphiquement (Figure 54). L'affinité est ensuite déterminée comme la moyenne des valeurs des K_D obtenus pour les 9 acides aminés présentant les CSP les plus significatives (Figures 55 et 56) : Ala42, Thr44, Gly46, Cys47, Leu116, Ser118, Ile119, Gly121, Thr147. Ces acides aminés sont situés dans l'environnement direct de Cys47 (acides aminés 42 à 46), mais aussi dans l'hélice Leu116 - Phe120 qui borde le site actif. L'affinité pour le CAT correspond à un $K_D = 4.5 \pm 0.6 \text{ mM}$ incluant une erreur sur les concentrations fixée arbitrairement à 20%, et le ΔG°_b est calculé par conversion (équation 5). En utilisant les règles de propagation de l'erreur, on obtient une énergie libre standard de liaison de $\Delta G^\circ_b = -3.2 \pm 0.1 \text{ kcal.mol}^{-1}$.

Dans le cas du 4MEC, la concentration maximale en ligand est d'environ 2 mM, soit un rapport de concentration final en $[L]/[P]$ de 6 avec $[P]_0 = 400 \mu\text{M}$. Au-delà, la limite de solubilité du ligand est atteinte, menant à une agrégation du ligand. Ce plus faible ratio a tout de même permis d'atteindre le début de la saturation de la protéine, et de déterminer une affinité de $K_D = 1.0 \pm 0.2 \text{ mM}$ soit $\Delta G^\circ_b = -4.1 \pm 0.1 \text{ kcal.mol}^{-1}$.¹⁸³ Ce résultat tient compte des 6 acides aminés suivants : Ala42, Thr44, Ser48, Thr50, Ile119 et Gly121, les seuls présentant une perturbation de déplacement chimique considérée comme significative (Figure 56). Il est probable que le nombre d'acides aminés présentant des CSPs non négligeables soit plus faible non pas parce que l'interaction est moins forte, mais parce que la concentration finale en 4MEC est encore un peu faible et certains acides aminés qui subissent l'interaction n'ont pas atteint la saturation, leur CSP est donc trop faible pour être prise en compte dans le cas présent.

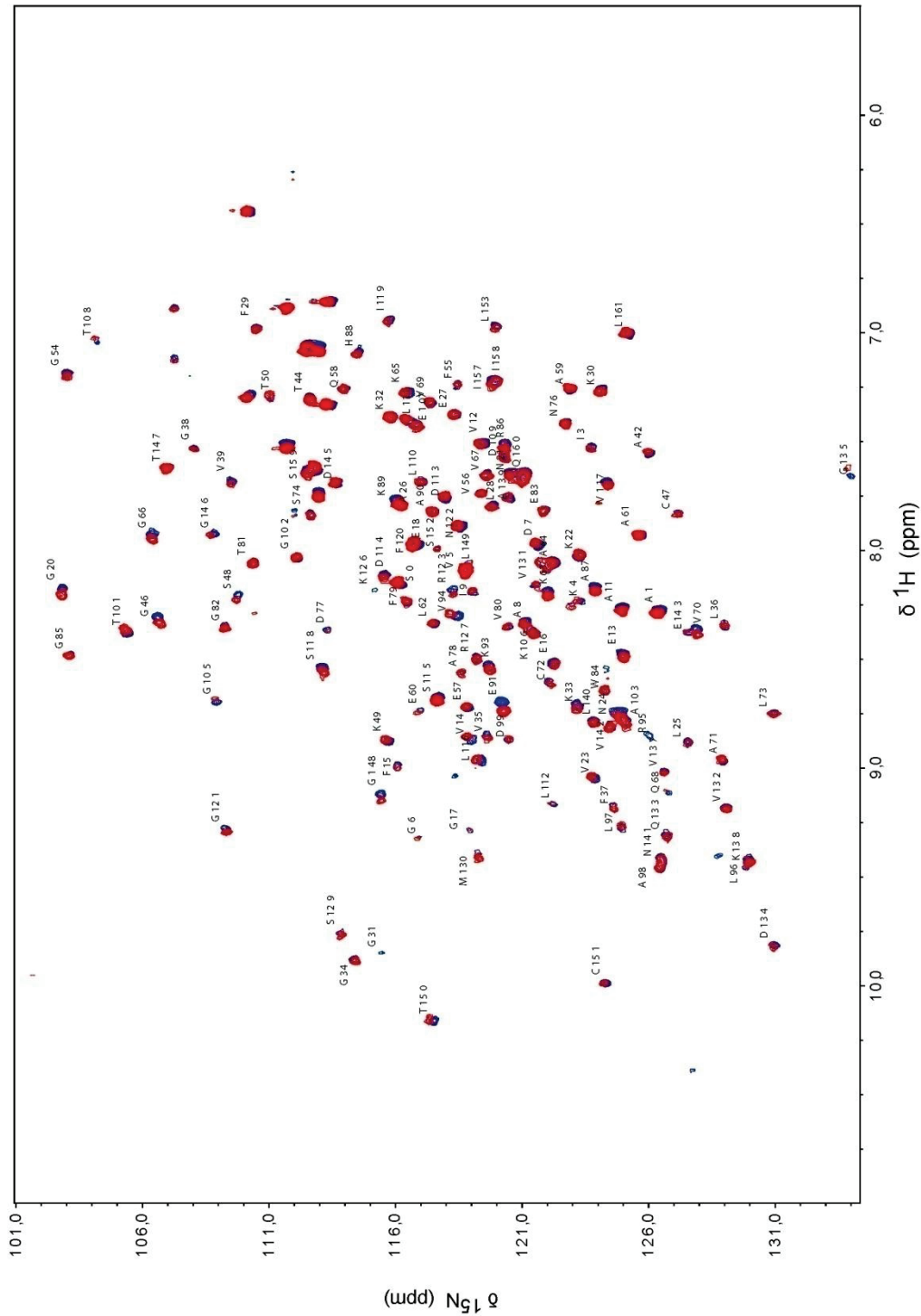


Figure 54. Influence du DMSO sur les signaux RMN de ^{15}N -hPRDX5 à 28°C, dans du PBS pH=7.4, enregistré à 600 MHz. Spectre ^1H - ^{15}N HSQC de hPRDX5 (200 μM) réduite par 2mM de TCEP, en absence de ligand (bleu), et en présence d'ajouts croissants de DMSO jusque 8% (rouge). Attribution des signaux des différents résidus de la protéine par le code à 1 lettre des acides aminés et numérotation selon leur place dans la séquence protéique. Variations considérées comme non significatives des signaux de la protéine jusqu'à atteindre 8% de DMSO.

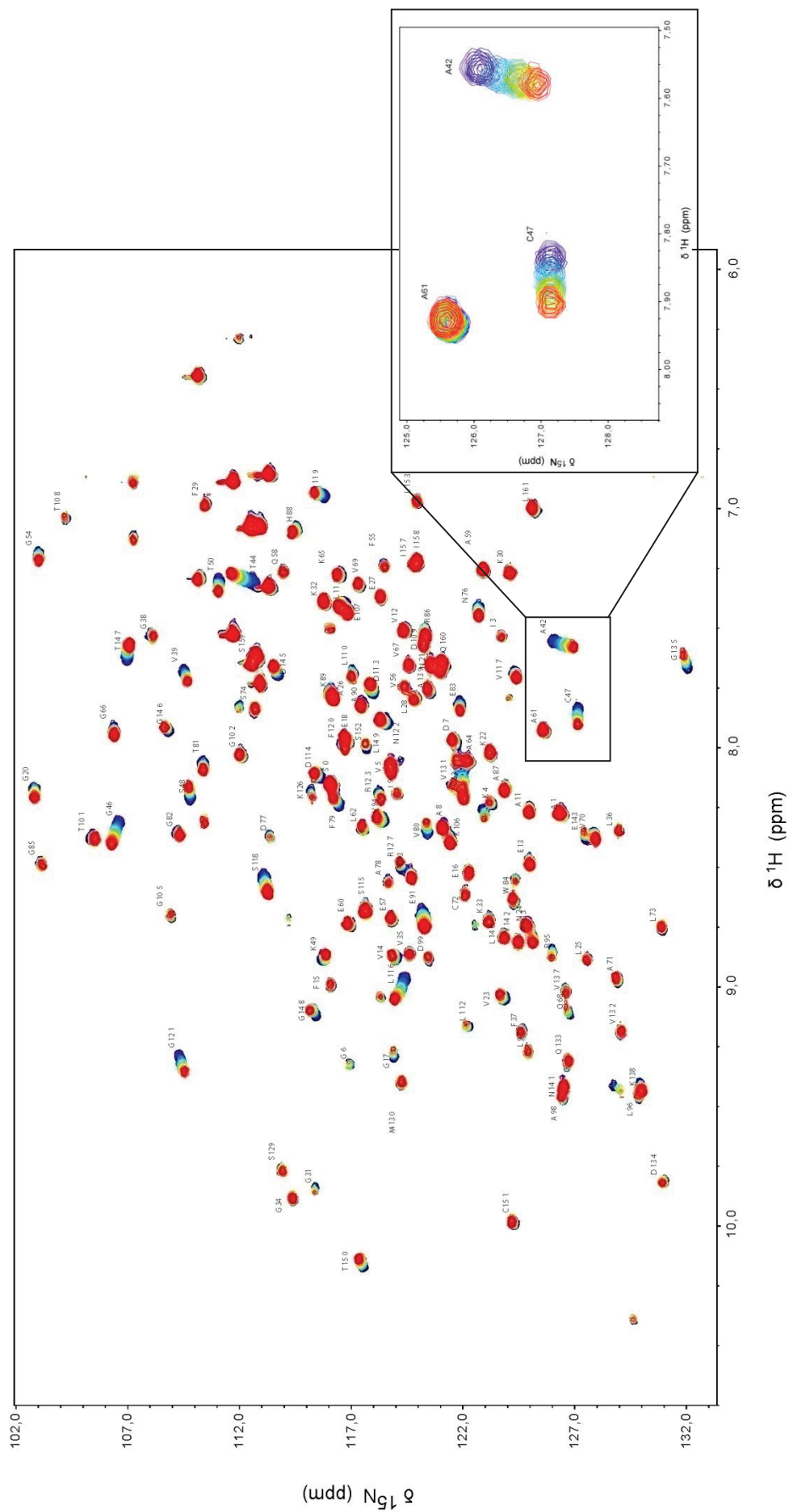


Figure 55. Spectre ^1H - ^{15}N HSQC de hPRDX5 (400 μM , 600 MHz, 28°C, pH=7.4) en présence d'ajouts croissants de CAT. Visualisation des CSPs induites par le ligand. Protéine réduite par 2 mM de TCEP. Spectre en absence de ligand (bleu), et en présence d'ajouts croissants de CAT jusqu'à atteindre la saturation à 10 mM (rouge). Attribution des différents signaux de la protéine par le code à 1 lettre des acides aminés, et numérotation selon leur place dans la séquence protéique.

Notons que les valeurs de K_D obtenues corrélient avec la valeur de K_D précédemment publiée par le laboratoire.⁹¹ En effet, il avait été déterminé pour CAT un K_D de 3.3 ± 0.5 mM avec hPRDX5, par une technique de RMN différente (STD voir paragraphe suivant).

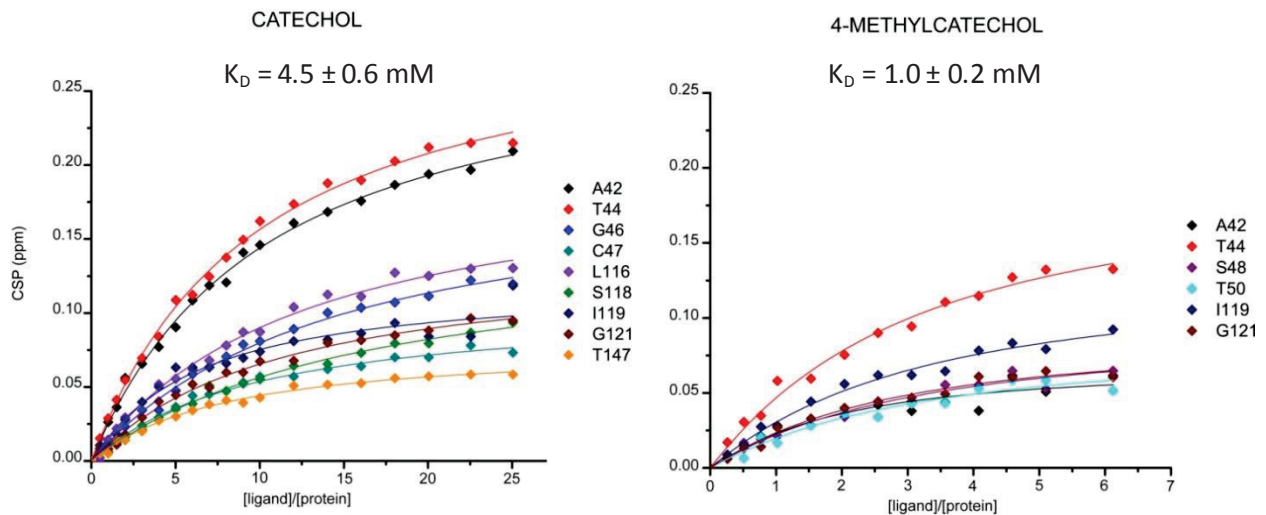


Figure 56. Courbes isothermes représentant la quantité de CSP induite par CAT et 4MEC sur les signaux ^1H - ^{15}N de hPRDX5 à 28°C et pH=7.4 dans un tampon PBS.

CSP (ppm) induite sur hPRDX5 par les ligands CAT et 4MEC en fonction de la quantité de ligand ajoutée exprimée en tant que rapport $[L]/[P]$, pour les acides aminés indiqués par leur place dans la séquence de la protéine. Concentration en protéine de 400 μM .

I. 3. 2 Etude des modes de liaison

A ce stade, il est intéressant d'étudier aussi l'affinité par une autre méthode couramment utilisée : la différence de transfert de saturation (STD). L'expérience STD permet d'évaluer la différence de saturation transférée depuis la protéine vers les différents protons du ligand. En principe cette saturation peut-être transférée différemment aux protons du ligand en fonction des conformations du complexe PL.

Dans le cas de CAT, la symétrie axiale du ligand ne permet pas de différencier le proton H_3 du proton H_6 ni le proton H_4 du proton H_5 qui ont des déplacements chimiques équivalents deux à deux (Figure 57). La titration effectuée par STD a donné lieu à une implication équivalente de ces deux groupes de protons tandis que pour le 4MEC, nous avons une différenciation des protons H_3 , H_5 et H_6 . Le proton H_3 est le plus saturé ($f_{STD} = 100\%$) tandis que les protons H_5 et H_6 dont le facteur STD maximal est autour de 60% ont subi un transfert saturation moins important (Figure 57). Ces résultats pourraient-ils laisser présager un mode de liaison prédominant ? Une cartographie d'épitopes effectuée par STD et publiée durant ce travail en 2014 montre une implication différente des protons aromatiques du

4MEC avec hPRDX5, où les protons H₅ et H₆ subissent la plus grande saturation.¹⁸² Or ces résultats s'appuient sur l'analyse d'une seule expérience effectuée en présence de 600 μM de ligand, et de DTT (1 mM). Comme nous avons vu que le DTT oxydé avait potentiellement la possibilité de se lier dans le site actif de hPRDX5, cela peut conduire à des disparités en terme d'affinité, mais aussi en terme de modes de liaison si le DTT est présent dans l'environnement de Cys47. Notons aussi que les affinités déduites diffèrent sensiblement d'un facteur 3 dans ce travail.¹⁸²

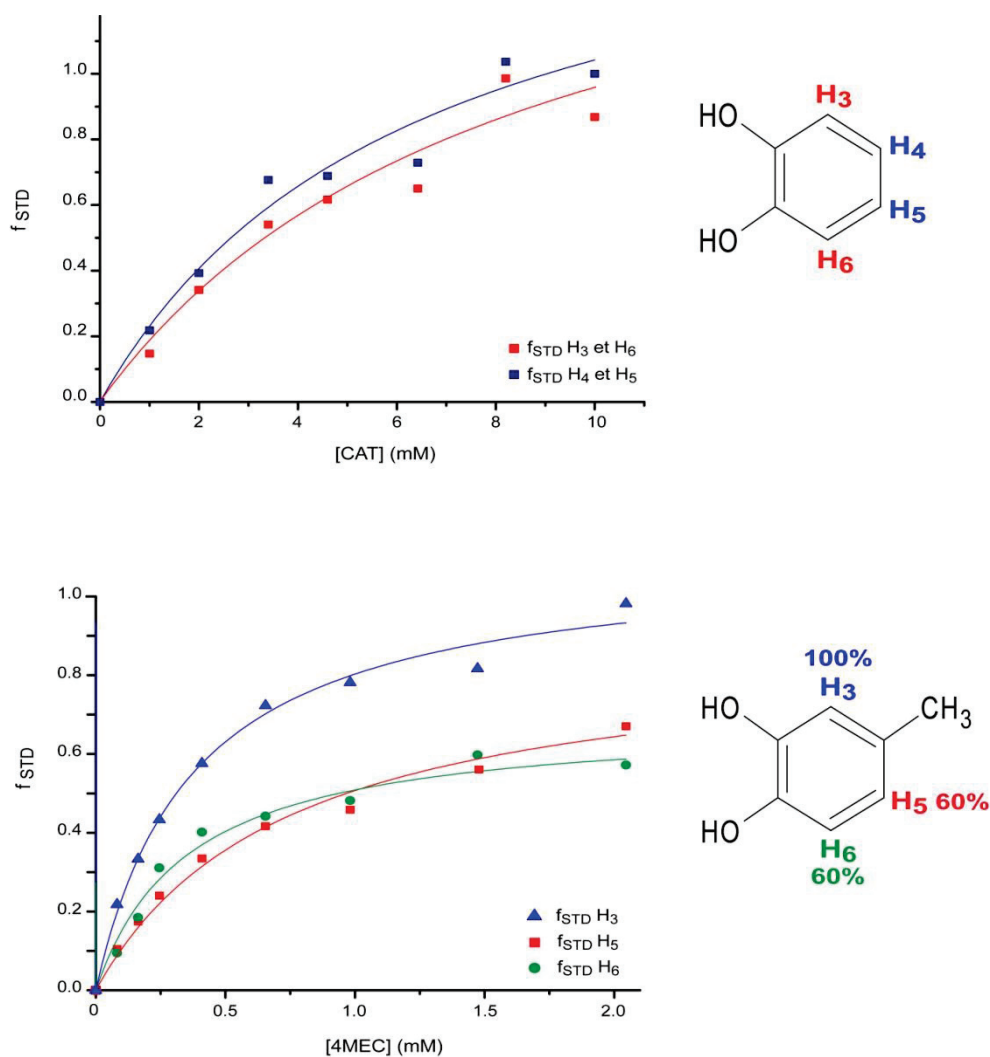


Figure 57. Cartographies d'épitopes obtenues par RMN STD pour les ligands CAT et 4MEC en interaction avec hPRDX5, à 28°C dans du tampon PBS, enregistré à 600 MHz.

Résultats des expériences par titration STD pour les ligands CAT (haut) et 4MEC (bas). Valeur de f_{STD} normalisée, tracée en fonction de la concentration en ligand (mM). Point expérimentaux mis en corrélation avec la courbe théorique (trait plein) obtenue par une équation du type hyperbole.

Si l'on compare maintenant ces résultats aux clusters conformationnels obtenus en FM, on peut rapprocher les conformations observées par transfert de saturation. Pour le CAT, nous avons vu par STD que l'implication des protons était équivalente, que ce soit par RMN ou par dynamique moléculaire. Dans le cas du 4MEC, les clusters donnent les 4 conformations discutées précédemment (Figure 52). Plus précisément, il y a une forte prédominance de la conformation B où le ligand se lie à SG Cys47 par le biais de 2 liaisons hydrogènes, avec le proton H₃ orienté vers l'intérieur du site actif tandis que les protons H₅ et H₆ sont orientés vers le solvant. Au total, les conformations B et C où H₃ possède cette même orientation correspondent à plus de 60% des conformations de liaison, contre à peine 40% pour les conformations A et D.¹⁸³

Nous pouvons donc conclure que ces résultats corrélerent avec ceux obtenus par les expériences STD, en sachant qu'en RMN ce sont la moyenne des différentes conformations qui sont observées.

I. 3. 3 Origines de la différence d'affinité entre CAT et 4MEC

La base de l'interaction des ligands catéchols avec hPRDX5 se trouve dans la possibilité de former différentes liaisons H entre les groupes hydroxyles du ligand et le thiolate de Cys47. En effet nous avons vu précédemment que le pKa de la chaîne latérale de la cystéine qui est théoriquement de 8.3, est abaissé autour de 5-6 dans le cas des PRDXs dû au résidu arginine présent dans l'environnement de la Cp. Ainsi, plus de 90% de la cystéine catalytique se trouve sous forme thiolate à pH 7 (Hall et al. et références citées).¹⁰⁶ Dans le cas de 4MEC, les interactions de vdW sont renforcées grâce à des interactions hydrophobes latérales avec les acides aminés Leu116, Ile119 et Phe120 de la chaîne A et Phe79 de la chaîne B, qui ont peut-être une conséquence sur l'amélioration de l'affinité.

Afin de comparer les CSPs induites par chacun des deux ligands sur hPRDX5, nous avons effectué des spectres ¹⁵N-HSQC de hPRDX5 en présence d'un ratio $[L]/[P]$ correspondant à $2 K_D$ pour chacun des deux ligands considérés, équivalent à environ 67% de la saturation totale. En représentant ces CSPs sur un histogramme empilé, nous pouvons voir que malgré une meilleure affinité, le 4MEC n'induit pas nécessairement une perturbation plus importante des déplacements chimiques des résidus de la protéine (Figure 58). La perturbation globale est certes améliorée car plus d'acides aminés sont impactés par l'interaction PL, d'où une affinité apparente plus importante. Cependant, comme montré en Figure 56, les acides aminés principalement impliqués dans l'interaction subissent une perturbation moins importante que dans le cas du CAT.

De façon hypothétique, on pourrait y voir une corrélation avec les résultats de FM. En effet nous avons vu que le 4MEC pouvait être échantillonné en 4 conformations de liaison différentes (Figure 53), contre 2 seulement pour le CAT.

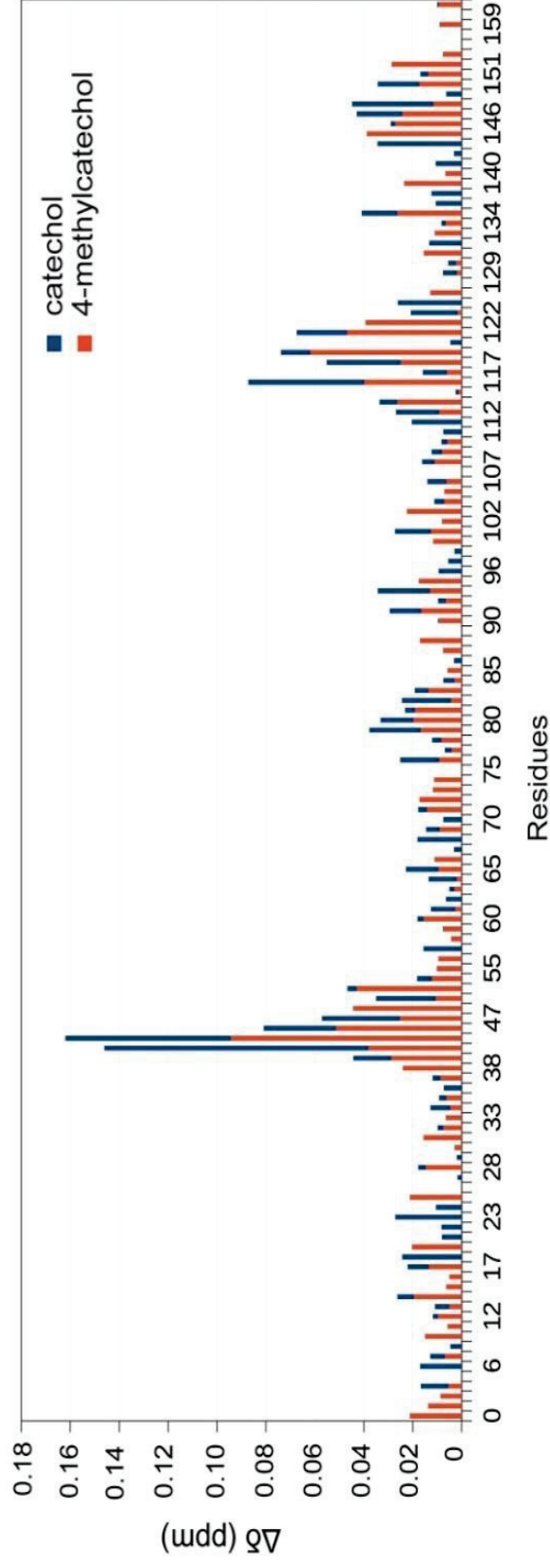


Figure 58. Histogramme cumulé représentant les CSPs induites par CAT et 4MEC sur les signaux RMN de hPRDX5, pour un rapport [L]/[P] correspondant à 2K_p. CSPs induites par CAT (bleu) et 4MEC (rouge) sur les signaux de hPRDX5, pour un rapport [L]/[P] correspondant à environ 67% de la saturation. Valeur de CSP correspondant à la somme des CSPs induites par CAT et 4MEC, et visualisation de la contribution de chacun des ligands. hPRDX5 200 μ M, [CAT] \approx 8 mM, [MEC] \approx 1.9 mM. Les données non représentées sur l'histogramme sont dues à des recouvrements spectraux ou aux acides aminés non observés en HSQC (résidus Proline).

Cela se traduit par une affinité moyenne plus importante mais qui semble moins sélective en conformation, résultat d'un moyennage plus important des CSP, reflet de la moyenne de toutes les conformations en échange rapide à l'échelle de temps du déplacement chimique RMN.

I. 3. 4 Comparaison des affinités calculées et expérimentales

L'intérêt principal d'étudier l'interaction de CAT et 4MEC avec hPRDX5 qui sont deux ligands possédant une affinité connue moyenne pour hPRDX5 était d'établir une corrélation entre les résultats obtenus par FM et ceux obtenus expérimentalement par RMN en solution. Le Tableau 6 ci-dessous reprend les valeurs d'affinité calculées par simulation puis celles déterminées par RMN. Si l'on compare les résultats d'affinité obtenus pour chacun des deux ligands, en termes d'énergie libre absolue de liaison ou de constante de dissociation, on peut voir que les valeurs expérimentales sont identiques aux valeurs calculées, à la fluctuation moyenne près.¹⁸³

Tableau 6. Affinités déterminées pour les ligands CAT et 4MEC pour hPRDX5 par FM et RMN.

Bilan des affinités calculées par FM et RMN en solution pour les complexes hPRDX5-CAT et hPRDX5-4MEC.

		CAT	4MEC
ΔG_b (kcal.mol ⁻¹)	FM	- 3.1 ± 0.2	- 4.2 ± 0.3
	RMN	- 3.2 ± 0.1	- 4.2 ± 0.1
K_D (mM)	FM	6.1 ± 1.8	0.9 ± 0.4
	RMN	4.5 ± 0.6	1.0 ± 0.2

Ces résultats étant, plusieurs approximations sont effectuées entre FM et RMN, qui pourraient tout de même mener à des écarts avec la réalité. Tout d'abord de la présence de DMSO et d'agent réducteur (TCEP) dans les expériences RMN ne sont pas prises en compte dans les calculs effectués par FM. Nous avons cependant vu que ni l'un ni l'autre n'avaient d'influence en RMN. Il existe aussi les approximations des champ de forces utilisés en FM (AMBER99SBildn).¹⁸⁶ Enfin, le PBS pH 7.4 (ions Na⁺, Cl⁻, K⁺, PO₄³⁻) utilisé en RMN est différent du milieu aqueux salin (NaCl) défini en FM (150 mM).

Grâce à l'étude des ligands CAT et 4MEC, nous pouvons conclure que l'estimation de l'affinité PL effectuée par le biais de simulations est en très bon accord avec les affinités expérimentales obtenues par RMN, dans une gamme d'affinité moyenne ($K_D \approx$ mM). Des études récentes ont montré que l'utilisation de la MD et la FM pour la détermination d'affinité de systèmes allant du micro au nano molaire était possible.^{72; 187} De plus les outils de *clustering* de trajectoire disponibles en aval de ces simulations permettent d'avoir une analyse de la prédominance de certaines conformations du complexe, ce qui est un atout pour la compréhension des mécanismes d'interaction. L'utilisation en

métadynamique d'un biais, et plus spécifiquement du potentiel externe en forme d'entonnoir utilisé en FM nous permet d'avoir un échantillonnage plus rapide et plus important qu'en dynamique moléculaire classique (NVT). En effet, les deux états essentiels à la description de l'interaction : l'état libre et l'état lié, sont en échange permanent ce qui permet une meilleure description des différentes conformations de liaison, même de faible énergie. Ces simulations introduisent une description quantitative de l'interaction PL, qui pourrait rendre obsolète toute autre approche empirique de *docking* moléculaire dans lesquelles le caractère intrinsèquement dynamique des protéines est ignoré. Nous pensons que cette approche de couplage entre l'échantillonnage de la métadynamique et des données apportées par RMN, ou avec toutes autres données expérimentales, est la meilleure voie pour interpréter de manière rationnelle les interactions entre deux systèmes en solution. Cette approche pourrait aussi permettre d'améliorer considérablement la compréhension des interactions PL, et la conception de ligands à haute affinité à partir de molécules *hits* simples avec une interaction faible, mais néanmoins sélective.

II. OPTIMISATION DE LIGANDS GUIDÉE PAR DYNAMIQUE MOLECULAIRE

La dynamique moléculaire fournissant une représentation plus réaliste et quantitative que les approches empiriques courantes (docking et scoring), on peut s'interroger sur la possibilité de se baser dessus pour la conception de ligands à haute affinité à partir de molécules *hits* simples. La deuxième partie du travail a été d'approfondir les observations obtenues par FM afin de déduire des faits qui pourraient permettre d'améliorer l'affinité du ligand pour hPRDX5. L'influence sur l'affinité de certaines modifications structurales du ligand ont été étudiées, et les interactions ayant lieu au sein du site actif entre le ligand et la protéine ont tenté d'être améliorées.

II. 1. Variations structurales et homologues du catéchol

Dans un premier temps, nous avons simulé par FM l'interaction de plusieurs molécules fragments structurellement proches du motif catéchol avec hPRDX5 (Figure 59).

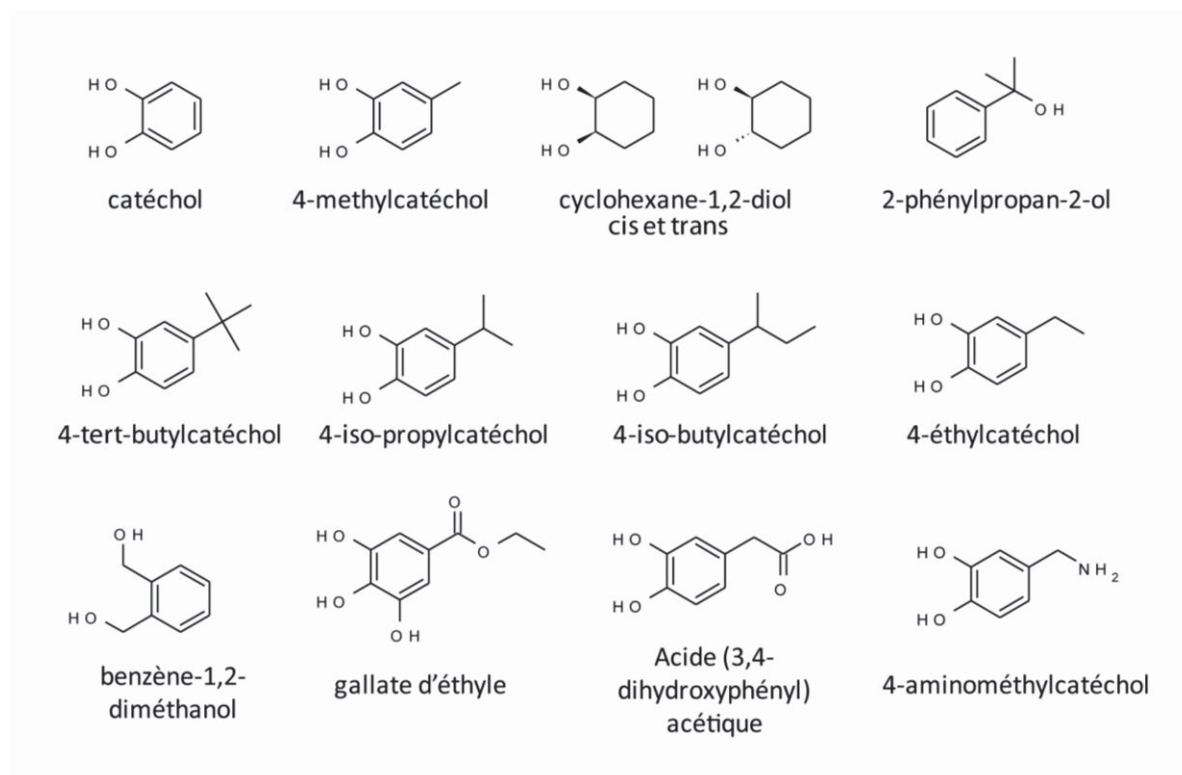


Figure 59. Structures des molécules fragments dont l'affinité pour hPRDX5 a pu être déterminée par FM, pour l'étude de l'interaction PL.

Structures des fragments étudiés, et leur dénomination correspondante.

L'étude de différents homologues nous a permis de comparer des caractéristiques structurales impliquées dans l'interaction. Parmi elles, l'aromaticité du cycle portant les deux groupements hydroxyles, leur position vicinale, ou la présence d'un groupement alkyle permettant des interactions de vdW supplémentaires à la périphérie du site catalytique de hPRDX5. Leur interaction avec hPRDX5 a été étudiée par des simulations de FM de durée variable, allant de 150 ns à 600 ns en fonction de la convergence du calcul. Leur énergie libre de liaison a été calculée, et les différents modes de liaison des fragments présentant les meilleures affinités pour hPRDX5 ont été étudiés. Les résultats d'affinité obtenus par FM pour ces ligands sont regroupés dans le Tableau 7 ci-dessous.

Tableau 7. Tableau récapitulatif des énergies de liaison ΔG_b° et constantes de dissociation K_D obtenues pour les différents ligands étudiés avec hPRDX5, par FM

Ligand	Temps de simulation (ns)	Dernières X ns pour calcul affinité	ΔG_b° (kcal/mol)	K_D (mM)
catéchol	500	100	- 3.0 ± 0.2	6.9 ± 2.1
4-méthylcatéchol	500	100	- 4.2 ± 0.3	0.9 ± 0.4
2-phénylpropan-2-ol	500	100	- 1.9 ± 0.3	36 ± 9
benzène-1,2-diméthanol	300	50	- 2.3 ± 0.1	22 ± 6
gallate d'éthyle	600	100	- 5.6 ± 1.1	0.3 ± 0.5
cyclohexanediol- <i>cis</i>	300	150	- 2.8 ± 0.5	12 ± 9
cyclohexanediol- <i>trans</i>	500	100	- 2.9 ± 0.5	12 ± 11
acide (3,4-dihydroxyphényl) acétique	500	150	- 5.8 ± 1.1	0.3 ± 0.9
4-aminométhylcatéchol	500	150	- 2.1 ± 0.6	40 ± 44
4-éthylcatéchol	500	150	- 3.6 ± 0.7	3 ± 2
4- <i>iso</i> -propylcatéchol	600	120	- 6.6 ± 1.0	0.02 ± 0.02
4- <i>iso</i> -butylcatéchol	500	150	- 4.3 ± 0.7	0.8 ± 0.6
4- <i>tert</i> -butylcatéchol	500	100	- 5.5 ± 0.9	0.11 ± 0.09

ΔG_b° = Valeur de l'énergie libre standard de liaison pondérée, prise sur les X dernières ns de simulation.

K_D = Valeur de la constante de dissociation du complexe PL pondérée au cours du temps, prise sur les X dernières ns de simulation

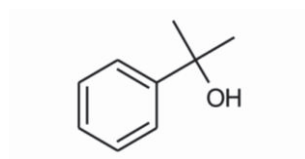
II. 1. 1 Influence des hydroxyles et de leur disposition relative sur le ligand

Nous avons vu en chapitre d'introduction que la réaction de réduction de H_2O_2 par hPRDX5 pouvait nécessiter une disposition particulière du peroxyde par rapport à la Cp. Dans ce modèle, le substrat H_2O_2 serait positionné dans le site actif, maintenu par un réseau de liaisons H, initiant la réaction par un mécanisme de type $\text{S}_\text{N}2$. Le thiolate Cp-S de hPRDX5 attaque un des groupes hydroxyles de H_2O_2 . L'état de transition de la réaction est atteint lorsque le groupe partant commence à se détacher alors que le thiolate se lie. Cet état est caractérisé par une distance O-O plus grande (de l'ordre de 1.75 Å au lieu de 1.4 Å), et au contraire une distance S-O plus faible (passage de 4.2 Å à 2.5 Å) selon Portillo-Ledesma et coll qui ont étudié le chemin réactionnel par des méthodes quantiques.¹⁵⁴ Ainsi, un ligand ayant deux groupements hydroxyles proches dans l'espace pourrait posséder une analogie structurale avec le peroxyde d'hydrogène et surtout, l'état de transition de la réaction enzymatique. Ils développeraient ainsi une affinité avec les hPRDXs, selon le fait que les analogues de l'état de transition ont en générale une bonne affinité pour le site récepteur des enzymes.¹⁵⁵ Dans un premier temps, l'influence de la position et du nombre de groupements hydroxyles sur l'affinité avec hPRDX5 a été étudiée. Dans le cas du catéchol la distance O-O est de 2.5 Å, mais elle est de 2.8-2.9 Å dans le cas des molécules de type cyclohexane *trans* et *cis* respectivement.

- Etude d'une molécule de type alcool

Nous avons centré notre étude sur des dérivés du catéchol, molécule aromatique possédant deux groupements hydroxyles en position ortho, dont l'affinité moyenne pour hPRDX5 a été montrée. Nous avons évoqué dans le chapitre introductif la possibilité que l'affinité du fragment catéchol vienne de l'agencement de ses deux groupements hydroxyles, permettant de mimer l'état de transition de la réaction catalytique de l'enzyme avec H_2O_2 . Quelle serait alors l'affinité pour hPRDX5 d'une molécule de type alcool ne possédant qu'un seul groupement hydroxyle ? Nous avons vu précédemment que la présence des deux groupements hydroxyles menait à différentes conformations de liaison dans le cas des ligands CAT et 4MEC. La présence d'un seul alcool pourrait alors amener soit un gain de stabilité de la forme liée en diminuant le nombre de conformations de liaison possibles, soit au contraire diminuer l'affinité globale de cette molécule pour la protéine suite à la perte du motif 1,2-diol. Pour répondre à cette question, l'interaction de hPRDX5 avec le phénol a tenté d'être modélisée, mais curieusement des problèmes de calculs instables ne nous ont pas permis d'effectuer la simulation. En parallèle nous avons étudié un autre mono-alcool, non phénolique cette fois, le phénylpropan-2-ol. Il s'agit du produit de la réaction des PRDX avec le peroxyde de cumène. Nous avons donc jugé utile d'estimer son affinité pour hPRDX5. D'autre part, cela pouvait également nous donner des informations quant à l'interaction avec hPRDX5 d'une

molécule possédant un noyau benzénique et une fonction alcool non phénol et de l'importance ou non que l'alcool soit sous forme phénolique.



Phénylpropan-2-ol

Sa simulation avec hPRDX5 donne lieu à 5 évènements d'association. La visualisation des différentes conformations de liaison nous montre que lorsque le groupement hydroxyle se lie à SG Cys47, le cycle aromatique tend à être orienté vers le solvant. L'analyse de la simulation a conduit à un $\Delta G_b^\circ = -1.9 \pm 0.3$ kcal/mol soit un $K_D = 36 \pm 9$ mM, moins bonne que celle du catéchol ($K_D = 4.0 \pm 0.5$ mM). La perte d'affinité peut s'expliquer soit par la perte du deuxième groupement hydroxyle qui permettait une conformation de liaison plus stable, soit par la position du groupement alcool qui n'est pas un phénol. La simulation avec le phénol n'ayant pas abouti, il était difficile de trancher ici quant à l'une de ces deux hypothèses.

- Etude d'une molécule de type 1,2,3-triphénol

Par l'étude des ligands CAT et 4MEC, nous avons vu que le ligand avait 2 conformations principales de liaison mettant en jeu 1 ou 2 liaisons H des groupements hydroxyles en position ortho. Nous nous sommes alors demandé quelle pouvait être l'affinité pour hPRDX5 d'une molécule 1,2,3-triphénol. Trois groupements hydroxyles adjacents permettent théoriquement la formation de 5 complexes PL différents (Figure 60), selon si l'interaction est faite via la formation d'une ou de deux liaisons H entre les groupements hydroxyles et le thiolate SG de Cys47. Le nombre de conformations théoriques de liaison possibles ont-elles un effet bénéfique sur l'affinité PL ou au contraire un effet délétère? Pourrait-t-on observer une prédominance marquée de l'une des conformations de liaison? Pour tenter de répondre à cette question, nous avons étudié le gallate d'éthyle.⁸⁰

La simulation de 600 ns effectuée sur le système hPRDX5-GAL a conduit à une affinité de $\Delta G_b^\circ = -5.6 \pm 1.1$ kcal.mol⁻¹ soit un $K_D = 0.3 \pm 0.5$ mM. Cette simulation a été poursuivie jusqu'à atteindre 600 ns afin d'avoir une meilleure convergence de l'énergie. En effet, un évènement de liaison entre 400 ns et 500 ns avait induit une forte fluctuation dans l'énergie libre de liaison moyenne. L'affinité moyenne converge vers une valeur stable en fin de simulation, nous pouvons conclure à une meilleure affinité pour hPRDX5 du GAL par rapport au 4MEC.

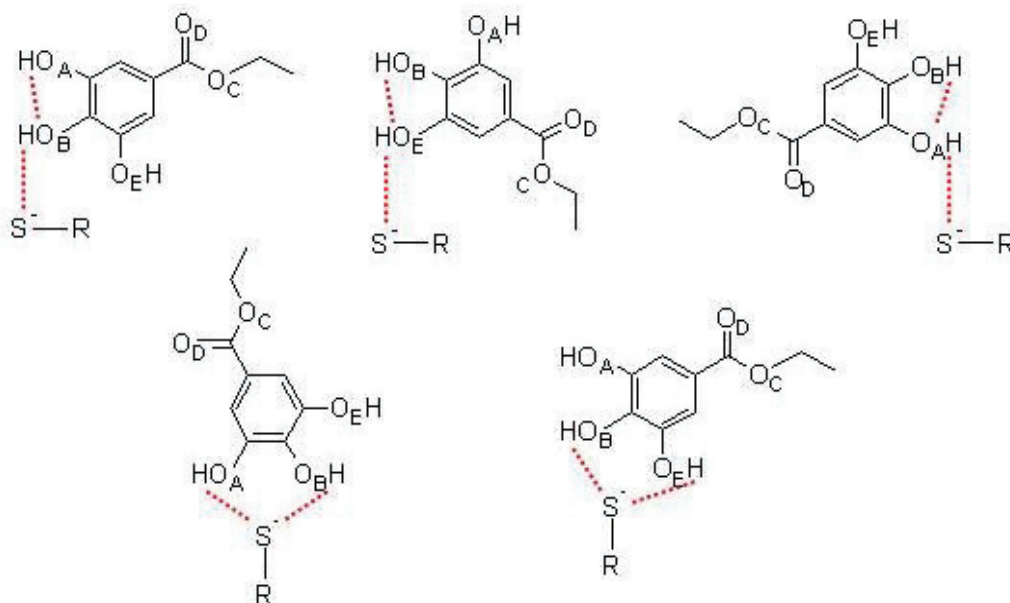


Figure 60. Cinq modes de liaison possibles du ligand gallate d'éthyle avec le thiolate Cys47.

3 modes de liaisons sont possibles entre GAL et SG Cys47 par le biais d'une seule liaison H. 2 autres modes de liaison sont possibles par le biais de deux liaisons H. La molécule ne possédant aucun centre chiral, uniquement ces 5 modes de liaison sont observables. Ces conformations sont aussi observables dans le cas de la liaison de GAL avec le carboxylate de Glu18 (chaîne B). R= Cys47

Cependant, la visualisation des surfaces d'énergie libre (FES) du complexe au cours du temps de simulation nous indique que l'ajout d'un groupement hydroxyle ne permet pas de stabiliser le complexe, qui va être en échange entre ses différentes conformations de liaison observées (Figure 59). Sur la FES (Figure 61), on peut voir que le bassin à 13 Å sur l'axe Z présente une énergie faible pour des valeurs d'angles variées en plus des deux minima énergétiques correspondant aux conformations de liaison observées pour CAT et 4MEC. Le duster prédominant (A) montre effectivement le ligand GAL en liaison avec Cys47 par le biais de 2 liaisons H formées par les groupements hydroxyles O_BH et O_EH . Si l'on regarde l'orientation de la molécule dans cette conformation, les groupements hydroxyles O_B et O_A sont alors alignés formant un angle $SG-O_A-O_B$ de 0° que l'on retrouve sur la FES (Figures 60 et 61).

Le cluster orange (B) met en évidence le même mode de liaison, mais avec une orientation différente du GAL dans le site actif. L'ajout de ce troisième groupement hydroxyle amène aussi une interaction périphérique avec Glu18 (chaîne B) caractérisée par le bassin situé à 19 Å sur l'axe Z, via la liaison de un ou plusieurs des groupements hydroxyles (Figure 61). En effet le résidu Glu18 (chaîne B) présente dans sa chaîne latérale un groupement sous forme carboxylate à pH = 7.4. Il est alors possible pour le ligand GAL de former une ou plusieurs liaisons hydrogènes avec ce résidu qui se trouve en surface de la hPRDX5, fortement accessible en comparaison de Cys47 qui se trouve au cœur de la cavité

hydrophobe du site actif. Les 3 autres clusters (C, D et E) rendent compte de cette interaction formée entre le ligand GAL et Glu18 (chaîne B) sous forme glutamate, via les groupements hydroxyles O_AH et O_BH dans plusieurs orientations. Ces 5 modes de liaisons qui ont été les plus représentés lors de la simulation ne mettent pas en évidence de conformation où le ligand GAL est lié à hPRDX5 formant un angle $SG-O_A-O_B$ de 60° . Cependant la FES montre que cette conformation à tout de même été échantillonnée. La valeur d'angle $SG-O_A-O_B$ ne suffit donc plus à décrire complètement le système et l'interaction PL du fait de la présence du troisième groupement hydroxyle et des nombreux modes de liaisons rendus possibles.

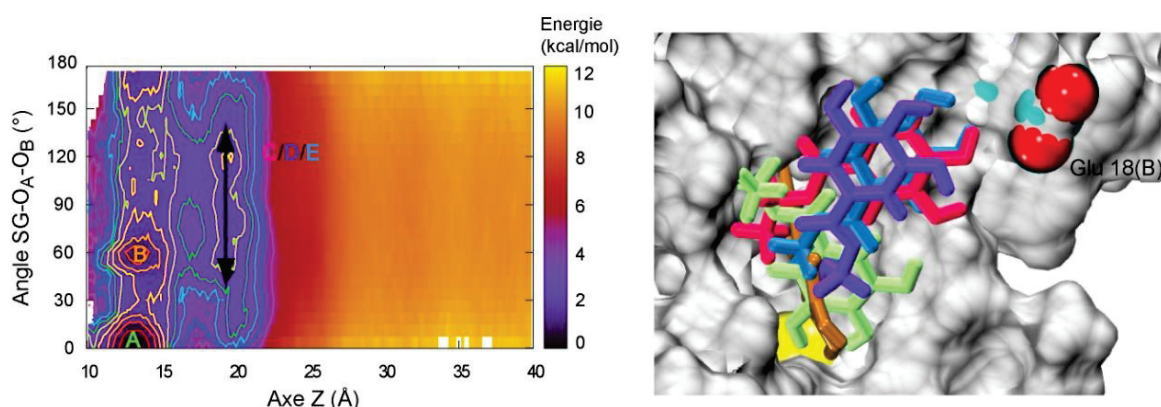


Figure 61. Analyse des modes de liaison du ligand gallate d'éthyle avec hPRDX5 par FM.

Gauche : FES 2D représentant l'énergie du complexe hPRDX5-GAL en fonction de Z (Å) et de l'angle $SG-O_A-O_B$. Courbes de niveaux représentées en lignes pleines avec un intervalle de $0.5 \text{ kcal.mol}^{-1}$. Droite : 5 clusters représentatifs des conformations de liaison du GAL pour hPRDX5 durant la simulation de FM. A : vert 50 % ; B : orange 12 % ; C/D/E : rouge violet et bleu respectivement 12 %, 10% et 16 % . Glu18(B) représenté par des sphères de vdW, colorées selon le type d'atome (bleu : carbone et rouge : oxygène).

Ce résultat a rendu nécessaire l'étude expérimentale du GAL avec hPRDX5 par RMN. Les titrations par 1H - ^{15}N -HSQC ont été effectuées par ajout de concentrations croissantes de GAL jusqu'à atteindre un ratio en concentration $[P]/[L]$ de 36, et une proportion de DMSO de 7% maximum. Les CSP ont été analysées en fonction de la concentration en ligand (rapport $[L]/[P]$) et comme dans le cas du CAT, les CSPs induites par le GAL sur hPRDX5 sont bien plus importants au niveau du site actif que dans le reste de la séquence (Figure 62). Les acides aminés pris en compte dans le calcul de la constante de dissociation du GAL pour hPRDX5 sont Ala42, Thr44, Gly46, Cys47 et Ser48 (au dessus du seuil indiqué en Figure 62). Ils sont tous positionnés dans le voisinage direct de la cystéine catalytique. Contrairement à la liaison des ligands CAT et 4MEC, les acides aminés hydrophobes de l'hélice Leu116 - Phe120 ne sont pas affectés en CSP (Figure 62). Glu18 (chaîne B), observé en FM comme entrant en interaction avec GAL, et l'on peut voir qu'il subit une perturbation qui dépasse le seuil de CSP. Cependant cette perturbation reste bien moins importante que celle induite au niveau du site actif puisque seule la chaîne latérale prend part à l'interaction. Or en HSQC, c'est la perturbation du

groupement NH de la chaîne principale qui est observée, et qui dans le cas présent est plus éloignée de l'interaction que le carboxylate.

L'affinité calculée est de $K_D = 1.3 \pm 0.2$ mM soit $\Delta G^\circ_b = -4.0 \pm 0.1$ kcal.mol⁻¹. Le résultat obtenu par RMN est légèrement plus faible que celui trouvé en FM ($K_D = 0.3 \pm 0.5$ mM), mais on reste dans l'ordre de grandeur attendu. La valeur de l'affinité du GAL étant similaire à celle trouvée dans le cas du ligand 4MEC en RMN, nous avons déduit de cette étude que l'ajout d'un troisième groupement hydroxyle n'amenait pas d'amélioration importante de l'affinité PL.

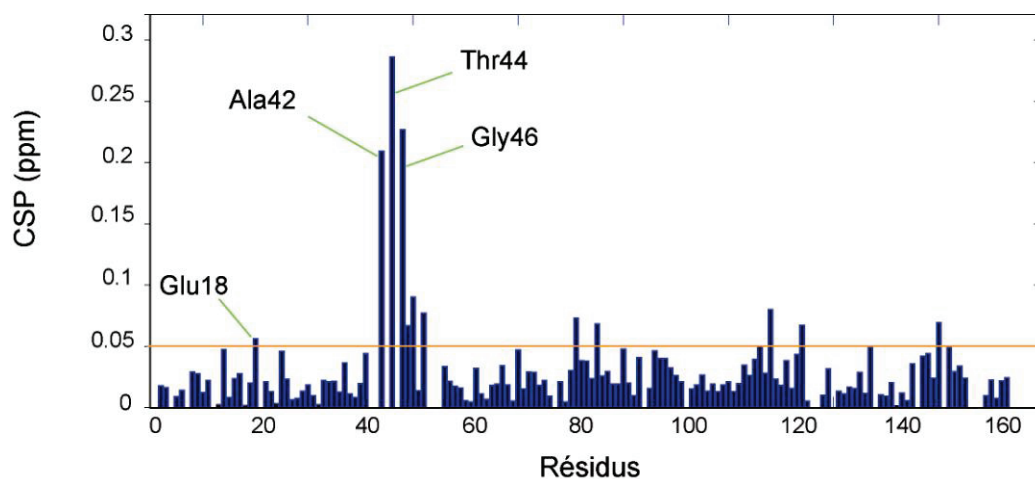


Figure 62. Histogramme représentant les CSPs induites par le ligand GAL (7mM) sur les signaux de ¹⁵N-hPRDX5 (200 μM), à 600 MHz, 28°C, pH=7.4.

Numérotation des résidus par leur place dans la séquence protéique. Valeur seuil des CSPs représentée par la ligne orange.

II. 1. 2 Influence de l'aromaticité

Le DTT oxydé, a été décrit ligand du site actif dans une structure cristalline de hPRDX5 (structure 3MNG.pdb).¹⁰⁶ Nous avons donc souhaité étudier l'influence de l'aromaticité d'un ligand sur son affinité pour hPRDX5 en étudiant une molécule analogue au DTT oxydé, le cyclohexane-1,2-diol. Cette molécule possède 2 groupements hydroxyles en positions vicinales 1,2 mais le caractère aliphatique de cette molécule induit que le cycle n'est pas plan, et donc que la distance entre les deux atomes d'oxygène d_{O-O} est plus grande (isomère *cis* = 2.9 Å ; isomère *trans* = 2.8 Å). Cette molécule possède deux isomères géométriques *cis* et *trans*, qui ont été étudiés par FM et RMN. L'isomère *trans* possède une plus grande analogie avec le DTT oxydé (DL) trouvé comme ligand ($d_{O-O} = 2.8$ Å), dans le cristal (Figure 63).¹⁰⁶ Le DTT est un réducteur, ce qui rend la mesure de son affinité pour hPRDX5 par RMN impossible puisque sa proportion sous forme réduite et oxydée en solution est difficile à contrôler, et en échange. L'étude des isomères du cyclohexane-1,2-diol permet donc, en s'affranchissant de ce problème d'oxydoréduction, d'estimer aussi l'affinité du DTT oxydé pour hPRDX5.

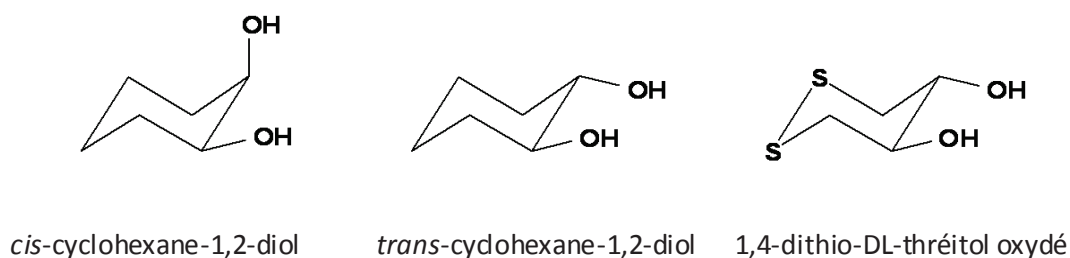


Figure 63. Structure des analogues *cis* et *trans* du cyclohexane-1,2-diol et du DTT sous sa forme oxydée cyclique.

Distances oxygène-oxygène pour ces 3 ligands : Isomère *cis* $d_{O-O} = 2.9 \text{ \AA}$; Isomère *trans* $d_{O-O} = 2.8 \text{ \AA}$; DTT oxydé $d_{O-O} = 2.8 \text{ \AA}$.

Les simulations de FM effectuées sur les deux isomères ont mené toutes deux à de très faibles affinités de l'ordre de 12 mM. Afin de déterminer expérimentalement l'affinité des isomères cyclohexane-1,2-diols pour hPRDX5 et déterminer les résidus affectés en déplacement chimique par l'interaction PL, nous avons étudié l'effet de ces deux ligands sur les signaux RMN de hPRDX5.

Dans le cas de l'isomère *cis*, des problèmes de solubilité ont mené à l'agrégation d'une partie de la protéine et du ligand en solution pour un rapport de concentration $[L]/[P] = 10$. Cette faible concentration n'a en conséquence pas permis une analyse fiable. En effet, on peut voir sur l'histogramme montrant les CSPs des résidus de la protéine, que la perturbation induite par l'ajout de 1,2-*cis*-cyclohexanediol mène à une perturbation moyenne inférieure au seuil de 0.05 ppm. Ce seuil ne permet pas de mettre en évidence les acides aminés perturbés par l'interaction (Figure 64). Uniquement deux acides aminés montrent des CSPs supérieures à la valeur seuil: Gly 46 et Ile 119. Les résidus du site actif ne montrent pas de perturbation plus importante que le reste de la protéine comme visible dans le cas des ligands CAT et 4MEC (partie I.3.3 du chapitre III).

Dans le cas de l'isomère *trans* nous avons pu effectuer une titration jusqu'à atteindre un ratio $[L]/[P]$ de 20 en solution. L'analyse des CSPs avec une valeur seuil choisie à 0.05 ppm mène à une mesure de $K_D = 20 \pm 2 \text{ mM}$ soit un $\Delta G_b^\circ = -2.3 \pm 0.1 \text{ kcal.mol}^{-1}$ (Figure 64).

Les valeurs obtenues nous montrent que si l'affinité est fortement réduite en absence du cycle aromatique, les acides aminés du site actif sont cependant significativement perturbés par le 1,2-*trans*-cyclohexanediol et non par l'isomère *cis*. L'interaction reste localisée au niveau du site actif, mais sa force est moindre en comparaison de celle formée avec le CAT, l'aromaticité est donc un paramètre important à prendre en compte dans la conception d'un ligand pour hPRDX5. Selon Hückel, cela se traduit par une molécule plane, possédant $4n+2$ électrons délocalisables dans un système cyclique (avec n le nombre de cycles) qui lui confèrent une stabilité particulière.

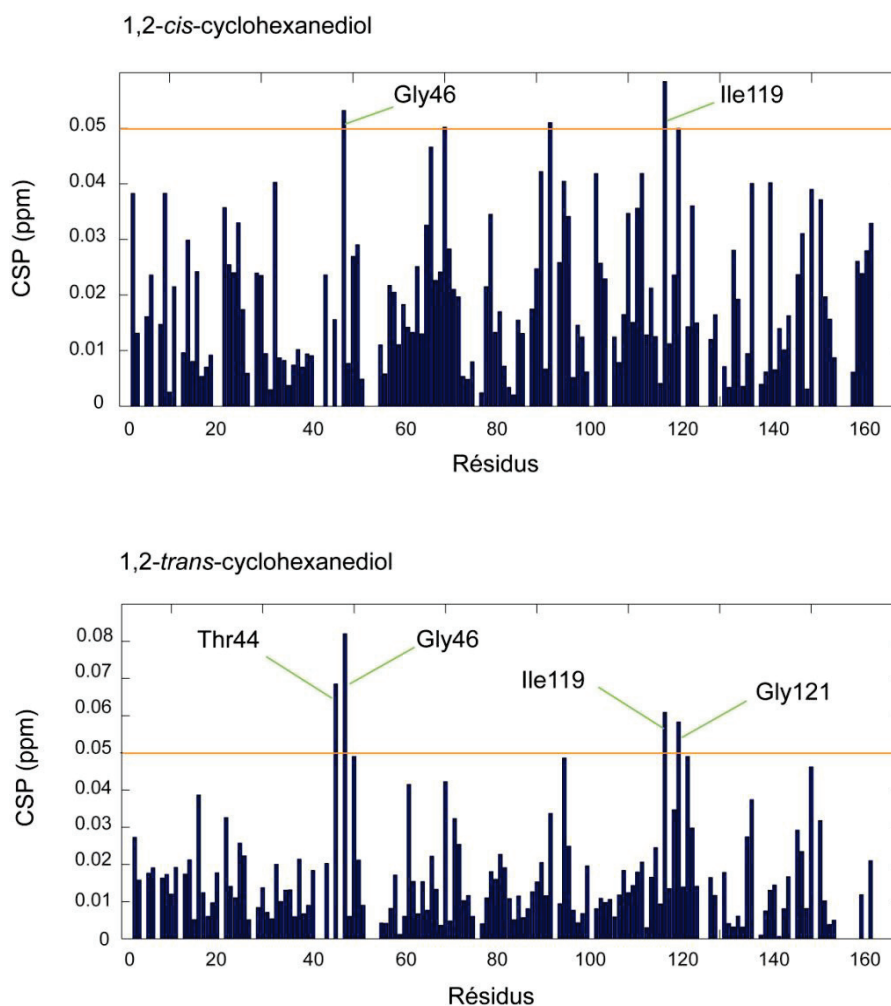


Figure 64. Histogrammes des CSPs induites par l'interaction des isomères *cis* et *trans* du cyclohexane-1,2-diol sur ^{15}N -hPRDX5, analysés par RMN à 600 MHz, 28°, pH=7.4.

CSPs induites sur les signaux RMN de hPRDX5 par l'ajout de 1,2-*cis*-cyclohexanediol (ratio [L]/[P] =10) et de 1,2-*trans*-cyclohexanediol (ratio [L]/[P] = 20). Valeur seuil de CSP représentée par la ligne orange.

II. 1. 3 Etude du substituant en position 4

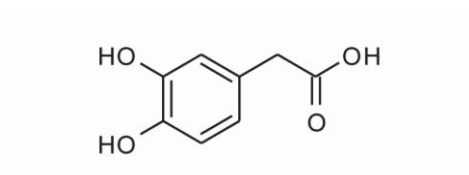
Nous avons vu en partie I.2 de ce chapitre de résultats (p.112) que le substituant alkyle placé en position 4 du cycle aromatique, formant le 4MEC, menait à une affinité améliorée pour hPRDX5 par rapport au CAT, du fait des interactions de vdW formées. Mais qu'en est-il lorsque l'on modifie la fonction chimique de ce substituant ? Les interactions éventuelles entre le ligand et la protéine vont être bien entendu différentes. Si dans le cas d'un groupement alkyle des interactions hydrophobes sont possibles, par remplacement de la fonction alkyle il va être possible d'induire d'autres interactions de type électrostatiques ou de liaison H selon la fonction chimique introduite. Nous avons déjà vu précédemment par l'étude du GAL que la fonction ester n'amenait aucune interaction bénéfique à la stabilisation du complexe, et menait à des interactions périphériques avec Glu 179 (chaîne B).

Nous avons souhaité estimer les affinités pour hPRDX5 de deux ligands (dérivés du catéchol) possédant deux fonctions chimiques courantes : la fonction acide carboxylique et la fonction amine.

- Introduction d'une fonction acide carboxylique

La fonction acide carboxylique est intéressante à étudier du fait de sa possible ionisation. A pH = 7.4, la fonction carboxyle est sous forme carboxylate, chargée négativement, ce qui permet de pouvoir créer des interactions de type coulombiennes avec des acides aminés chargés positivement, tel que l'arginine.

Nous avons vu en introduction que hPRDX5 possède une Arg conservée dans toutes les classes de hPRDXs, Arg127 dans le cas de hPRDX5. Or cet acide aminé posséderait un rôle dans la réaction de réduction catalysée par hPRDX5 en stabilisant la Cp par le biais de liaison H, et en lui donnant un proton afin de restaurer sa fonction acide sulfénique après libération de la molécule d'eau en fin de réaction. La molécule possédant une fonction carboxylique que nous avons testée est l'acide (3,4-dihydroxyphényl)acétique, dérivée du catéchol possédant une fonction acide en position 4. Son affinité pour hPRDX5 a été estimée par FM.



acide (3,4-dihydroxyphényl)acétique

Au cours de la simulation, 7 événements de liaison ont pu être observés, avec une alternance des conformations de liaison à une ou deux liaisons hydrogènes entre le ligand et SG Cys47 de hPRDX5. On obtient après analyse de la trajectoire une valeur d'énergie libre absolue de liaison de $\Delta G_b^\circ = -5.8 \pm 1.1 \text{ kcal.mol}^{-1}$ soit un $K_D = 0.3 \pm 0.9 \text{ mM}$.

Le *clustering* des conformations les plus fréquemment observées nous indique que ce ligand tend à se lier préférentiellement à l'acide aspartique Asp77 (chaîne B), localisé à l'interface avec la chaîne A de l'homodimère (Figure 65). Comme on peut le voir sur la Figure 64, Asp77 (chaîne B) est enfoui au cœur de la structure de la cavité, et c'est par ses groupements hydroxyles que le ligand vient se lier, par le biais de liaisons H. Cette interaction semble prédominante par rapport à l'interaction avec Cys47 : 72% de conformations en liaison avec Asp77 (chaîne B) contre 20% de conformations liées à SG Cys47 (Figure 65).

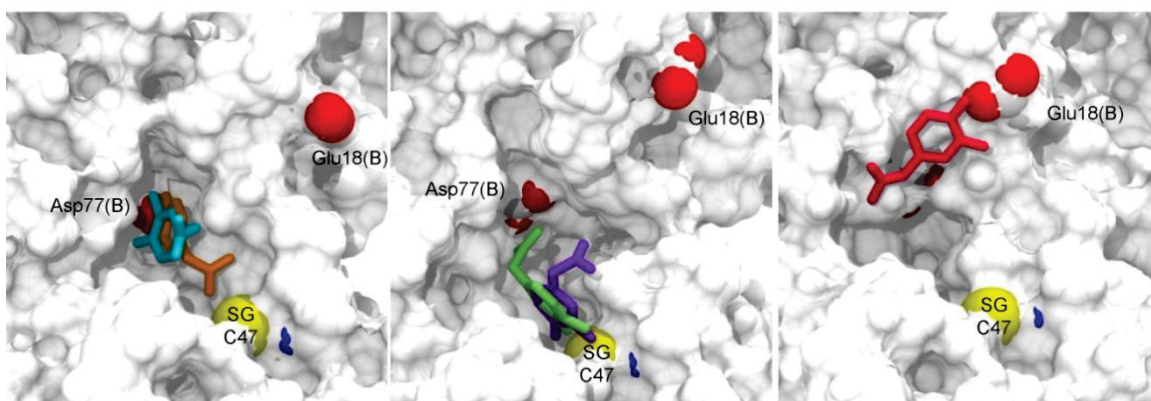


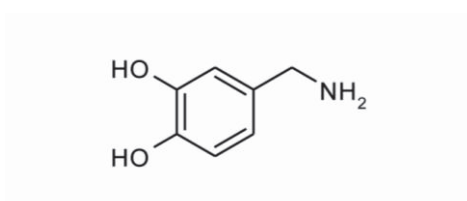
Figure 65. Clusters des conformations de liaison observés pour l'acide (3,4-dihydroxyphenyl) acétique avec hPRDX5 par FM.

Visualisation de la liaison avec Asp77 et Glu18 de l'autre monomère (B). SG Cys47, Glu18 (chaîne B) et Asp77 (chaîne B) sont représentés par des sphères de vdW. Le ligand est représenté sous forme de bâtons. Proportion des différentes conformations de liaison analysées : Orange : 46% ; Bleu : 26% ; Violet : 11% ; Vert : 9% ; Rouge : 8%.

L'interaction ionique recherchée par l'introduction d'une fonction acide n'est donc pas observée, et l'affinité mesurée lors de cette simulation ne correspond pas à l'interaction de l'acide (3,4-dihydroxyphenyl)acétique avec Cys47.

- Introduction d'une fonction amine

De la même manière que précédemment, l'intérêt d'introduire une fonction amine vient de sa capacité à former des interactions coulombiennes ou des liaisons H avec les acides aminés ionisables (Glu, Asp, Cys) de hPRDX5. Le 4(méthylamino)catéchol représente donc l'homologue cationique direct de la molécule précédente.



4-(méthylamino)catéchol

Le pKa d'une amine primaire étant de 10.66, le ligand se trouvera sous forme ionisée $-NH_3^+$ à pH = 7.4. Théoriquement ce ligand peut donc interagir avec des acides aminés chargés négativement, de type Glu, Asp ou encore Cys sous forme thiolate. Or nous avons vu précédemment que les acides aminés Glu18 et Asp77 de la chaîne B de l'homodimère pouvaient entrer en interaction avec le ligand induisant une délocalisation de l'interaction principale. Nous cherchons ici à faire interagir la partie ammonium du ligand avec ces acides aminés, afin que les groupements hydroxyles se lient prioritairement à Cys47.

Durant la simulation, 6 courtes périodes de liaisons ont été observées permettant un bon échantillonnage des conformations libres et liées. La simulation a donné lieu à une énergie moyenne convergeant vers une valeur stable de : $\Delta G_b^\circ = -2.1 \pm 0.6 \text{ kcal.mol}^{-1}$ soit $K_D = 40 \pm 44 \text{ mM}$ (Figure 66).

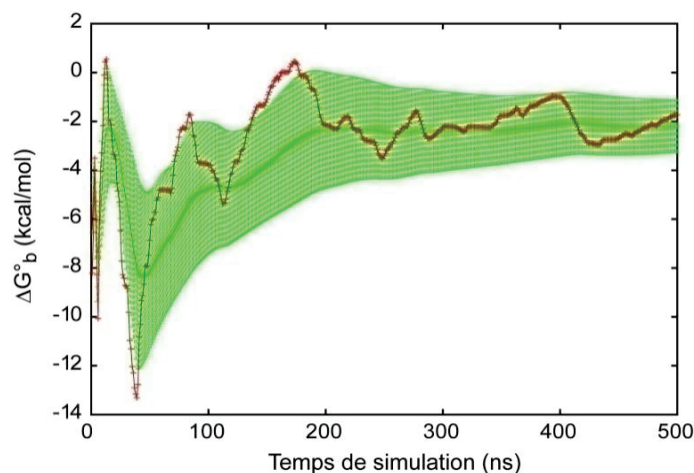


Figure 66. Visualisation de l'énergie libre de liaison du complexe 4-(méthylamino)catéchol-hPRDX5 au cours de la simulation FM.

ΔG_b° (kcal.mol^{-1}) en fonction du temps de simulation représentée en trait plein rouge. ΔG_b° moyenne pondérée au cours du temps en trait plein vert, et ses fluctuations en lignes pointillées.

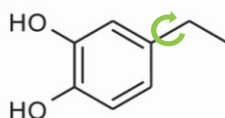
L'affinité obtenue pour ce complexe est moins bonne que celle obtenue dans le cas du catéchol seul ou d'autres dérivés car la fonction amine amène la formation d'interactions à la surface de la protéine, ne permettant pas la stabilisation de l'état lié du complexe PL. La modification de la fonction chimique du groupement en position 4 n'a donc pas permis de mettre en évidence de nouvelles interactions, et l'affinité de ces composés pour hPRDX5 est plus faible du fait de la diminution des interactions hydrophobes et des interactions secondaires observées comme avec Asp77 (chaîne B) par exemple. Nous avons alors choisi de travailler sur l'optimisation du groupement alkyle afin d'améliorer les interactions hydrophobes à l'intérieur du site actif.

II. 1. 4 Modification de la chaîne carbonée aliphatique

L'étude de différents groupements fonctionnels en position 4 n'ayant pas permis de mettre en évidence de nouvelles interactions PL stabilisantes, nous avons choisi de conserver un groupement alkyle pour les interactions de van der Waals qu'il peut former avec les acides aminés présents dans le voisinage du site actif tels que l'hélice Leu116 - Phe120. Afin d'améliorer ces interactions de vdW, nous avons étudié plusieurs ligands par FM en modifiant la longueur, le volume et le nombre de torsions libres de ce groupement alkyle.

- Etude du 4-éthylcatéchol par FM

Le premier ligand que nous avons étudié, est le 4-éthylcatéchol (4ECAT) possédant une chaîne alkyle à 2 carbones en position 4. Ayant observé une légère amélioration de l'affinité entre les ligands CAT et 4MEC due à l'interaction de vdW formée entre le groupement méthyle et l'hélice Leu116 - Phe120, nous avons formulé l'hypothèse que l'ajout d'un carbone supplémentaire permettrait d'augmenter cette interaction.



4-éthylcatéchol

La simulation FM de 4ECAT avec hPRDX5 a donné lieu à 8 évènements de liaison durant lesquels le ligand se lie alternativement à Cys47 par une seule liaison H ou deux liaisons H. La FES obtenue en fonction des variables collectives CV1 et CV3 montre un bassin avec une dispersion importante des conformations liées en termes d'angle $SG-O_A-O_B$ (entre 30 et 90°). L'affinité calculée de ce complexe PL est proche de celle obtenue dans le cas du CAT et du 4MEC : $\Delta G_b^\circ = -3.6 \pm 0.7 \text{ kcal.mol}^{-1}$ soit $K_D = 3 \pm 2 \text{ mM}$ (Figure 66). La dispersion des conformations associée au degré de liberté supplémentaire donné à la chaîne ne favorise pas, dans ce cas, la stabilisation d'une seule conformation. Cela mène à une affinité à la convergence, globalement équivalente à celle du 4MEC, donc légèrement meilleure que celle du CAT pour hPRDX5 (Figure 67).

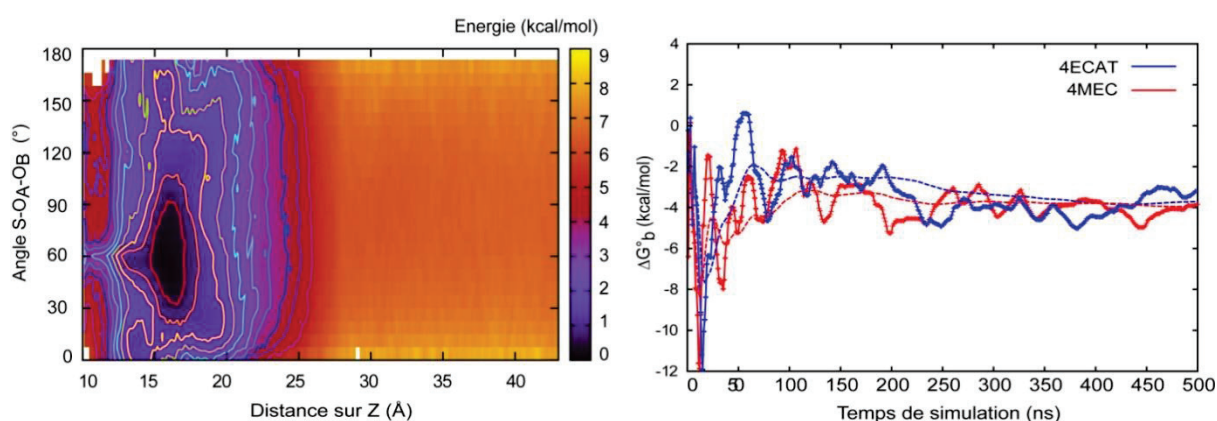


Figure 67. Analyse FM du complexe hPRDX5-4ECAT : FES CV1-CV3 et ΔG_b° .

Gauche) FES tracée en fonction des variables collectives CV1 (Z (Å)) et CV3 (angle $SG-O_A-O_B$ formé par l'atome de soufre de la cystéine catalytique, et les deux atomes d'oxygène O_A et O_B du ligand). Courbes de niveau en lignes pleines, toutes les 0.5 kcal/mol. Droite) Comparaison de ΔG_b° pour les ligands 4MEC et 4ECAT avec hPRDX5 au cours de la FM. Energie libre de liaison ΔG_b° (kcal/mol) en fonction du temps de simulation (ns) en ligne pleine, et l'énergie libre moyenne pondérée en fonction du temps en ligne pointillée.

- Etude des ligands 4-isopropylcatéchol et 4-isobutylcatéchol par FM

L'élongation seule de la chaîne carbonée linéaire n'ayant pas d'effet, nous avons étudié l'influence de la ramification de cette chaîne. L'augmentation se fait alors en termes de volume et non uniquement de longueur. Le but recherché est toujours d'améliorer les interactions de vdW formées par le ligand à l'intérieur du site actif afin d'aider à la stabilisation du complexe lié PL, et donc a de meilleures affinités.

Les deux ligands étudiés sont les suivants :



Ces 2 ligands ont la particularité d'être directement dérivés du catéchol par l'ajout en position 4 d'un groupement *iso*-propyl (3 carbones) et *iso*-butyl (4 carbones). La chaîne alkyle est donc dans les deux cas ramifiée et non plus linéaire, ce qui lui induit un volume plus important et plus de possibilités de liaisons de vdW. De plus, le ligand 4-*iso*-butylcatéchol possède un degré de liberté au niveau de son groupement alkyle. On remarque l'analogie de ces deux dérivés avec les acides aminés Val, Leu et Ile au niveau des chaînes latérales. Dans les protéines, ces acides aminés aliphatiques peuvent former des patches hydrophobes. Les simulations effectuées sur ces deux ligands en interaction avec hPRDX5 donnent lieu aux résultats d'affinité présentés dans le Tableau 8 ci-dessous.

Tableau 8. Résultats d'affinité des ligands 4-*iso*-propylcatéchol et 4-*iso*-butylcatéchol pour hPRDX5 estimés par FM

Ligand	Durée de la simulation (ns)	X ns prises en compte pour le calcul de ΔG_b°	ΔG_b° (kcal/mol)	K_D (mM)
4- <i>iso</i> -propylcatéchol	600	120	- 6.6 ± 1.0	0.02 ± 0.02
4- <i>iso</i> -butylcatéchol	500	150	- 4.3 ± 0.7	0.8 ± 0.6

Dans le cas du ligand 4-*iso*-propylcatéchol, la simulation a donné lieu à 7 événements de liaison de durée variable (de 1 ns à 80 ns). Durant les périodes de liaison les plus longues (60 et 80 ns), le ligand est lié à hPRDX5 par le biais d'une seule liaison hydrogène, tandis que lors des courtes périodes de liaison (< 20 ns), le mode de liaison mettant en jeu 2 liaisons H prédomine. La surface d'énergie libre

(FES) en fonction des variables collectives CV1 et CV3 donne un bassin d'énergie minimum bien délimité, centré sur Cys47 ($Z = 13 \text{ \AA}$) avec 2 valeurs d'angle correspondantes aux 2 modes de liaison observés précédemment. La différence d'énergie est d'environ 10 kcal.mol^{-1} avec le complexe libre en solution, laissant présager une bonne affinité de ce ligand. Un second bassin est observé entre 15 et 20 \AA sur Z , de plus faible énergie (Figure 68).

Dans le cas du ligand 4-*iso*-butylcatéchol, comportant un carbone et un degré de liberté supplémentaire au niveau du groupement alkyle, les évènements de liaisons observés ont été beaucoup plus rares. Seulement 2 évènements, durant lesquels la liaison avec la protéine s'est faite par le biais d'une seule liaison H. Le bassin d'énergie observé correspondant au complexe PL est moins bien défini dans le cas présent en termes d'angle $\text{SG-O}_A\text{-O}_B$, et la différence d'énergie entre l'état libre et l'état lié, d'environ 7 kcal.mol^{-1} , est plus faible que celle observée avec le dérivé *iso*-propyl (Figure 68). Cette observation est confirmée par la valeur de l'affinité qui est nettement meilleure dans le cas du ligand 4-*iso*-propylcatéchol. L'énergie des systèmes est à la convergence dans les deux cas, et la moyenne pondérée stable en fin de simulation.

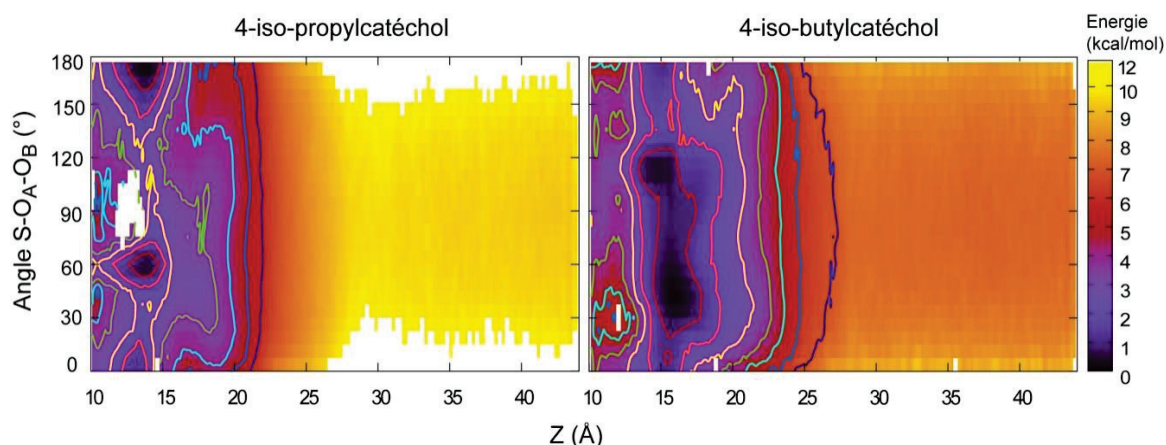


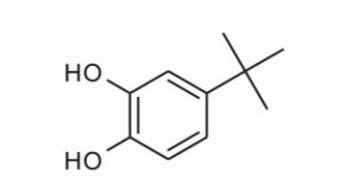
Figure 68. Différences en termes d'énergie de liaison, entre les complexes formés par les ligands 4-*iso*-propylcatéchol et 4-*iso*-butylcatéchol avec hPRDX5.

Surfaces énergétiques tracées en fonction des variables collectives CV1 (Z (\AA)) et CV3 (angle $\text{SG-O}_A\text{-O}_B$ formé par l'atome de soufre de la cystéine catalytique, et les deux atomes d'oxygène O_A et O_B du ligand). Courbes de niveau dessinées en traits pleins avec un pas de 1 kcal.mol^{-1} .

Le degré de liberté apporté au groupement aliphatique dans le cas du 4-*iso*-butylcatéchol apparaît ici défavorable en termes d'affinité. La dispersion conformationnelle mènerait alors à une différence d'énergie plus faible. Le ligand 4-*iso*-propylcatéchol par contre présente une affinité améliorée. Ce dérivé n'est pas commercialement disponible, mais l'homologue le plus proche, le dérivé *tert*-butylé, possédant un carbone supplémentaire, l'est et pourrait donc être testé expérimentalement aisément.

II. 1. 5 Etude de l'affinité du dérivé 4-*tert*-butylcatéchol

Avant d'effectuer la simulation par métadynamique, nous avons pris le temps de faire une simulation de dynamique NVT sur le 4TBC en utilisant le même protocole que pour les dérivés CAT et 4MEC. Au cours de cette simulation, nous avons pu observer la liaison d'une des molécules de 4TBC à l'un des sites actifs de hPRDX5 (Cys47 chaîne B), durant un temps de simulation approximant les 400 ns ce qui est un temps de résidence bien plus long que ceux observés pour les ligands précédemment étudiés par dynamique NVT : CAT et 4MEC (I.2.1 p.106), et une indication qualitative d'une meilleure affinité.



4-*tert*-butylcatéchol

- Estimation de l'affinité protéine – ligand par FM

L'analyse de la trajectoire FM sur le système hPRDX5-4TBC a permis de comptabiliser 6 périodes de liaison (Figure 69. B). La convergence a été vérifiée par le tracé du dépôt des potentiels au cours du temps de simulation (Figure 69.C), et la valeur de l'énergie libre de liaison de ce complexe a pu être calculée à cette convergence. On obtient une affinité correspondante de $\Delta G_b^\circ = -5.5 \pm 0.9 \text{ kcal.mol}^{-1}$ soit un $K_D = 0.11 \pm 0.09 \text{ mM}$.

Cette affinité est améliorée en comparaison de celle obtenue dans le cas de 4MEC. Cette amélioration pourrait provenir de meilleurs contacts de vdW entre le groupement *tert*-butyl et les acides aminés hydrophobes formant l'hélice Leu116 - Phe120, ou bien seulement de l'augmentation des contacts hydrophobes avec la surface de la protéine. En effet les contacts sont fortement augmentés du fait du nombre de groupes méthyles. Cependant, le gain d'affinité ne modifie en rien la dispersion de l'état lié, qui reste partagé entre différentes conformations de liaison.

En cartographiant l'énergie du complexe en fonction des deux variables collectives CV1 et CV3 définies précédemment (Figure 44), nous avons obtenu la surface d'énergie libre du complexe PL sur laquelle sont visibles les deux bassins d'interaction observés dans le cas du CAT et du 4MEC, dont l'énergie diffère seulement de $0.5 \text{ kcal.mol}^{-1}$ (Figure 70). Ces bassins représentent les deux conformations de liaison possibles du 4TBC avec hPRDX5, selon le nombre de liaisons H formées entre les groupements hydroxyles et SG Cys47 (Figures 46). Le *clustering* regroupe ces deux conformères, mais aussi une conformation supplémentaire correspondant à l'interaction du ligand

avec Glu18 (chaîne B) par 2 liaisons H formant un angle SG-O_A-O_B de 60° environ (Figure 70). Le centre de masse du ligand ayant été défini en prenant en compte tous les atomes lourds du ligand, il en résulte que le centre de masse du ligand en liaison avec Cys47 (Z = 13 Å) se trouve spatialement très proche de celui du ligand en liaison avec Glu18 (chaîne B) (Z = 16 Å).

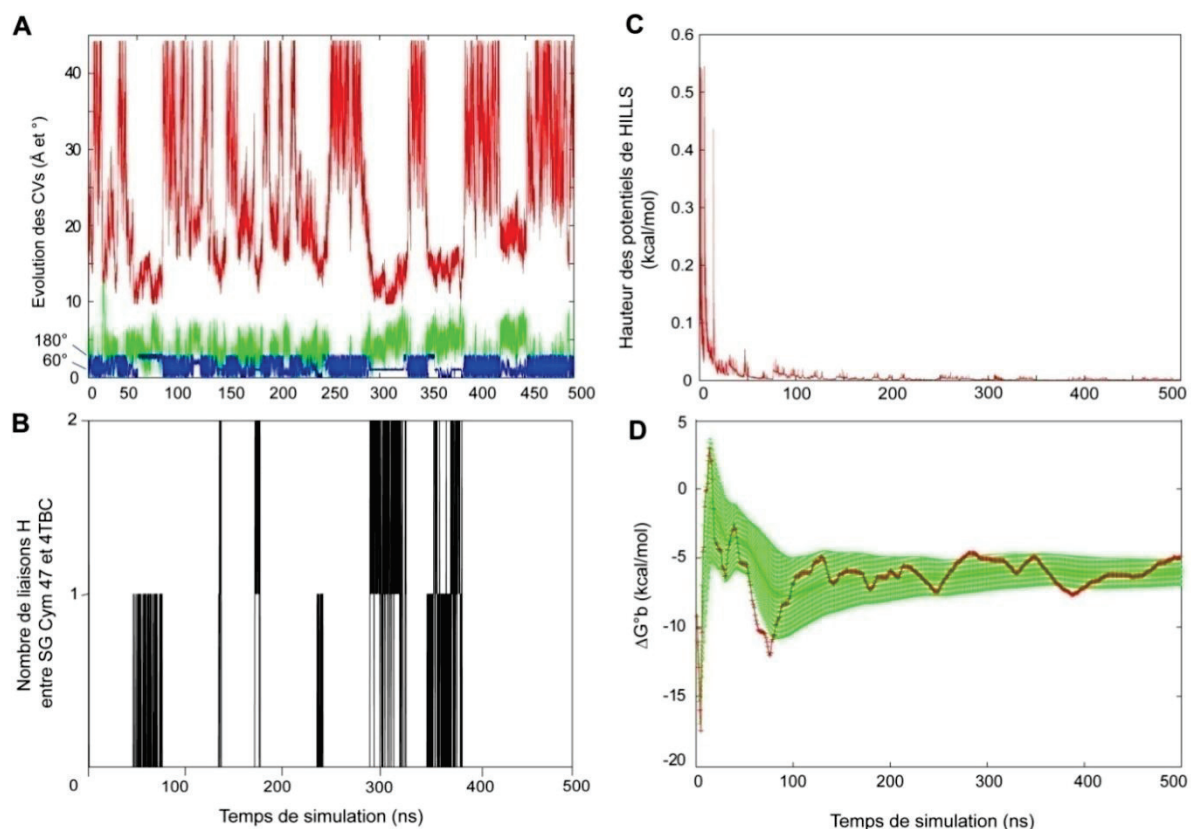


Figure 69. Analyse de la simulation du complexe 4TBC avec hPRDX5 par FM.

A) Evolution des CVs au cours de la simulation. Distance du ligand sur l'axe Z (rouge), distance par rapport à ce même axe (vert), angle formé par SG Cys47 et les deux atomes d'oxygène du ligand (bleu). B) Liaisons H formées entre 4NPC et SG Cys47 au cours du temps de simulation. C) Evolution du dépôt des potentiels de HILLS au cours de la simulation. D) Evolution de l'énergie libre de liaison du complexe PL durant la simulation (rouge), énergie moyenne pondérée au cours du temps (vert) et sa fluctuation.

La proportion de ces 3 conformations a pu être estimée, en dénombrant le nombre de conformations observées dans chacun des « clusters » au cours de la simulation. Il en résulte que la conformation où le ligand est en liaison avec Cys47 par liaison de ses 2 groupements hydroxyles est donnée comme majoritaire avec 50% de présence, contre 35% pour la forme liée à 1 liaison H dont l'angle SG-O_A-O_B formé est de 180°, et 15% pour la conformation en interaction avec Glu18 (chaîne B) (Figure 70). Une structure cristallographique est disponible à la PDB (entrée : 4K7O) et présente le ligand en liaison avec Cys47 selon la conformation à une seule liaison H présentant un angle SG-O_A-O_B de 180° (bassin A).¹⁸²

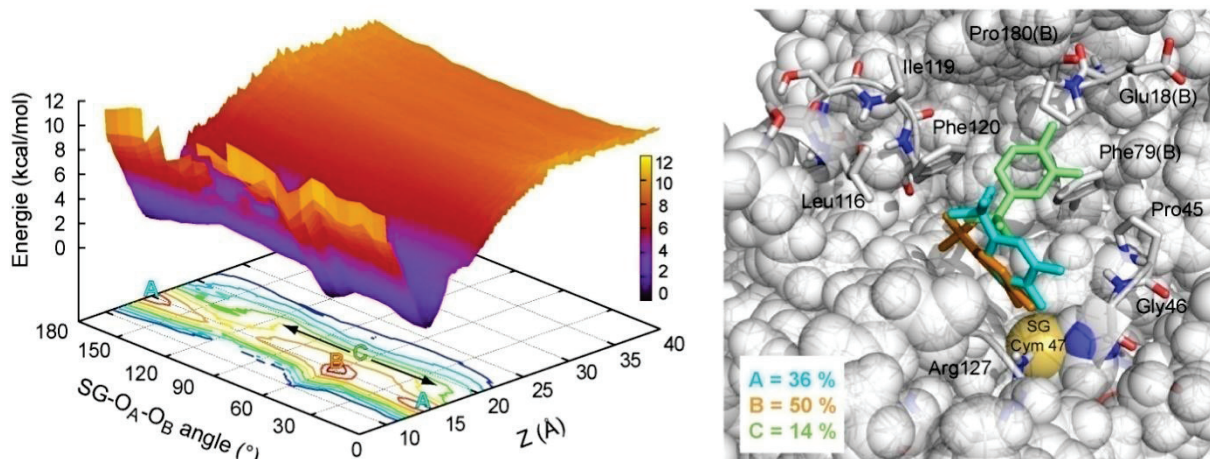


Figure 70. Analyse des 3 modes de liaison du ligand 4TBC avec hPRDX5 par FM.

Gauche : Surface d'énergie libre (FES) du complexe PL en fonction des variables collectives CV1 (Z (Å)) et CV3 (angle $SG-O_A-O_B$ formé par l'atome de soufre de la cystéine catalytique, et les deux atomes d'oxygène O_A et O_B du ligand). Courbes de niveau dessinées en traits pleins avec un pas de $0.5 \text{ kcal.mol}^{-1}$ puis de 1 kcal.mol^{-1} à partir de 1 kcal.mol^{-1} . Droite : Représentation des 3 conformations de liaison obtenues après « clustering » géométrique de la trajectoire en utilisant une distance seuil de 2 \AA , et représentatives des 2 bassins énergétiques. Cys47 est représentée par une sphère (van der Waals) jaune, et les conformations de liaisons colorées en accord avec leur appartenance aux différents bassins d'interactions (A : bleu 36% ; B : orange 50% ; C : vert 14%).

- Etude de l'interaction hPRDX5-4TBC par RMN en solution

L'affinité de 4TBC pour hPRDX5, comme pour les dérivés précédemment étudiés, a été validée expérimentalement par RMN en utilisant des expériences de type ^{15}N -HSQC. La gamme de concentration utilisée pour le ligand est comprise entre 0 et 3 mM correspondant à une quantité maximale de DMSO de 5% dans l'échantillon. Lors de l'analyse des déplacements chimiques des signaux de hPRDX5 en présence de 4TBC, 9 acides aminés ont montrés des CSP supérieures à 0.05 ppm : Ala42, Thr44, Gly46, Cys47, Ser48, Thr50, Glu83, Gly121, et Gly148 (Figure 71).

Comme pour les autres catéchols, les acides aminés perturbés sont au voisinage direct de la cystéine catalytique : Ala42, Thr44, Gly46, Ser48, Thr50, incluant la cystéine catalytique Cys47 elle-même. On remarque que les acides aminés de l'hélice Leu116 - Phe120 ne présentent pas de perturbation de déplacement chimique significative. En traçant la valeur des CSPs pour ces 9 acides aminés en fonction de la valeur du rapport des concentrations $[L]/[P]$, on obtient l'affinité suivante pour le complexe hPRDX5-4TBC : $K_D = 190 \pm 40 \mu\text{M}$ soit un $\Delta G^\circ_b = -5.1 \pm 0.1 \text{ kcal.mol}^{-1}$.

Nous pouvons conclure que les valeurs sont en corrélation avec celles obtenues par FM, dans la marge d'incertitude déterminée. Ces résultats portant sur l'interaction du 4TBC avec hPRDX5 ont été publiés en 2016 dans le journal *Biochemistry*.¹⁸⁸

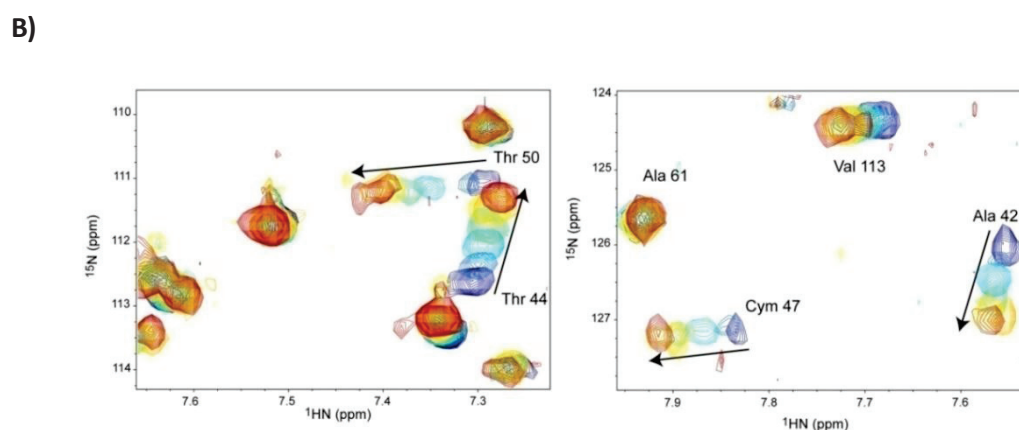
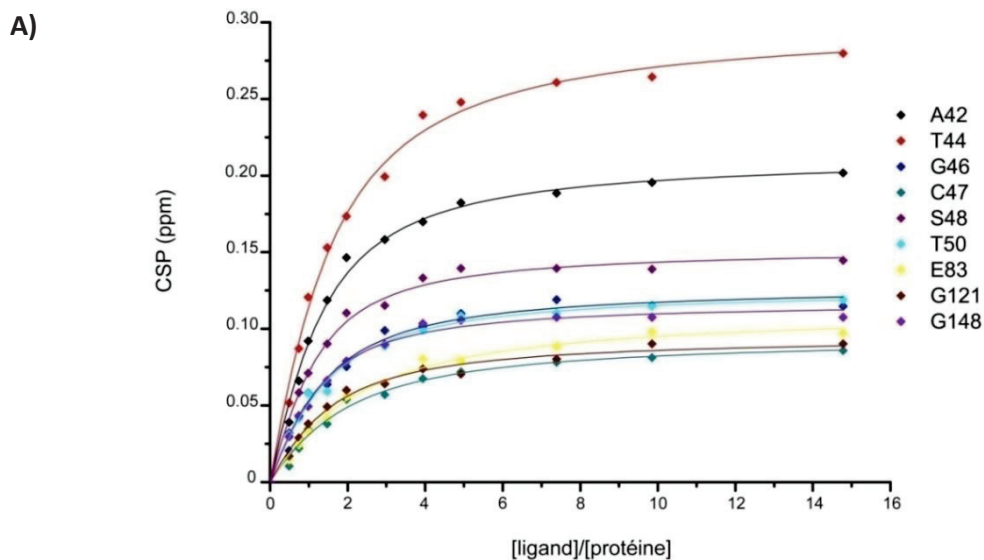


Figure 71. Interaction du 4TBC avec ^{15}N -hPRDX5 (200 μM) étudiée par RMN ^{15}N -HSQC, à 600 MHz, 28°C, pH=7.4.

A) Isothermes de liaison des acides aminés impliqués dans l'interaction protéine-ligand. Courbes représentant la perturbation de déplacement chimique en ppm, en fonction de la quantité de ligand ajoutée exprimée comme le rapport des concentrations de ligand sur protéine. Point expérimentaux sous forme de losange, isothermes de liaison théorique associées, pour un modèle 1:1, en courbe pleine. B) 2 expansions de spectres montrant l'évolution de déplacement chimique des acides aminés Ser48, Thr44, Ala42, et Cys47 lors des expériences HSQC effectuées avec des ratios de concentrations [4TBC] sur [hPRDX5] allant de 0 (bleu) à 15 (rouge).

- Cartographies d'épitopes par expériences STD

En FM, on se base sur la proportion relative des différentes conformations de liaison observées, afin de mettre en évidence un ou plusieurs modes de liaison prédominants. L'expérience RMN STD permet elle aussi de détecter une éventuelle conformation majoritaire de liaison par le biais de la quantité de saturation reçue par les différents protons du ligand, au cours de l'interaction. Nous avons souhaité étudier l'interaction du 4TBC par STD afin de détecter une possible corrélation avec le mode de liaison prédominant proposé en FM.

La titration par STD a été effectuée en présence de 10 μM de protéine jusqu'à atteindre une concentration en 4TBC de 2 mM. Le calcul des facteurs d'amplification STD f_{STD} (relation **35** p.94) pour chaque concentration en 4TBC permet d'obtenir les isothermes montrés en Figure 72. Il ressort que, contrairement aux observations faites dans le cas du 4MEC, H₅ et H₆ sont les protons les plus saturés à hauteur de 100% contre 60% pour H₃. Si l'on compare ces résultats à ceux obtenus lors du *clustering* géométrique effectué lors de la simulation FM du ligand 4TBC avec hPRDX5, on ne trouve pas de corrélation directe. En effet, la conformation principale du complexe observée en FM liant SG Cys47 à 4TBC par 2 liaisons H (50%) montre H₃ tourné vers l'intérieur du site actif, et H₅/H₆ orientés vers le solvant (Figure 70) qui devraient donc avoir un f_{STD} moins élevé. Nous discuterons plus loin ce fait.

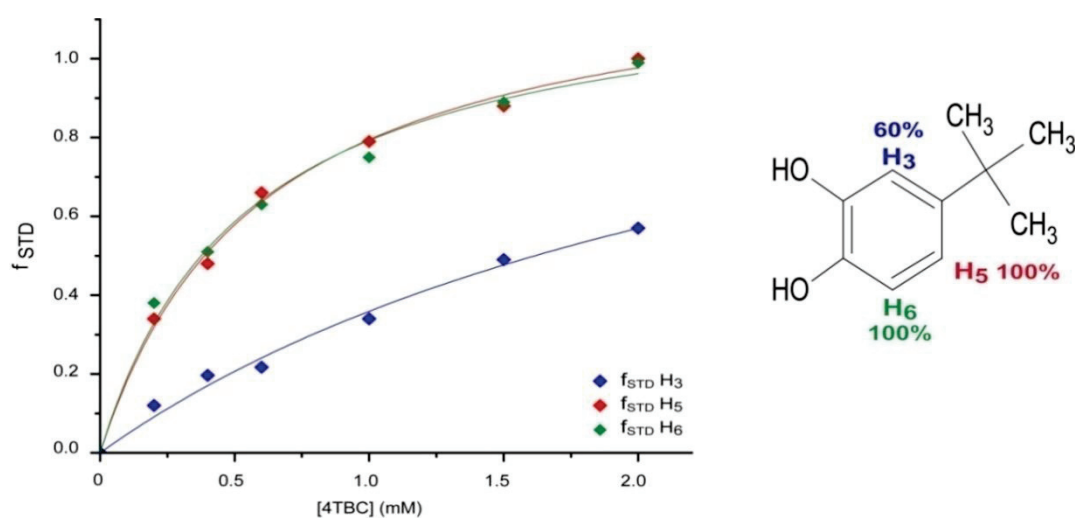


Figure 72. Implication des différents protons aromatiques du 4TBC dans son interaction avec hPRDX5 (10 μM) par expérience STD, à 600 MHz, 28°C, pH=7.4.

Courbes représentant le f_{STD} pour chacun des protons aromatiques de 4TBC en fonction de la concentration en ligand. Correspondance des points expérimentaux avec une courbe théorique (traits pleins) calculée par une équation de type hyperbole $y = A x / B + x$.

- Conclusion sur l'étude de l'affinité du 4TBC

Nous avons pu voir avec l'étude du 4TBC qu'il y a une amélioration de l'affinité par rapport aux ligands CAT et 4MEC mais qui ne vient pas d'une amélioration des interactions de vdW entre le groupement *tert*-butyl et les chaînes latérales de l'hélice Leu116 – Phe120 spécifique de hPRDX5 par rapport aux autres isoformes humaines des PRDXs où elle est absente. Nous avons obtenu un bon agrément entre simulation et expérience. L'analyse RMN nous montre que les CSPs les plus importantes concernent bien les acides aminés au voisinage de Cys47 comme dans le cas de CAT et 4MEC mais que le patch hydrophobe Leu116 – Ile119 – Phe120 ne subit pas de CSP significative. On peut imaginer que la molécule ne s'adapte pas de manière optimale à la géométrie du site actif

impliquant que le ligand n'arrive pas à rester en contact permanent avec la cystéine et le patch hydrophobe en même temps, ce qui empêche la stabilisation de l'état lié du complexe PL.

Nous avons cherché à renforcer les différentes interactions pouvant avoir lieu dans le site actif, entre le ligand et hPRDX5, tout en donnant un caractère plus flexible au ligand, ce qui nous a mené à la conception du ligand 4-*néo*-pentylcatéchol (4NPC).

II. 2. Conception et étude du 4-*néo*-pentylcatéchol

II. 2. 1 Conception du 4-*néo*-pentylcatéchol

Comme nous l'avons vu en introduction, le site actif de hPRDX5 forme une cavité d'accès direct depuis le solvant. En prenant en compte les acides aminés qui forment ce site actif (Figure 31 p.70) majoritairement hydrophobe, plusieurs interactions sont possibles lors de la formation du complexe :

- Bien entendu la liaison hydrogène entre donneur et accepteur de protons, qui est la base de l'interaction des molécules *o*-catéchol avec hPRDX5, et plus précisément avec le thiolate de Cys47 située au cœur du site catalytique.
- Des interactions hydrophobes avec les acides aminés de l'hélice Leu116 - Phe120 ainsi qu'avec Pro19 (chaîne B) et Phe79 (chaîne B) situées dans le fond de la cavité, vers l'intérieur de la structure.
- Le π -*stacking* entre les ligands possédant un motif catéchol, et les acides aminés aromatiques présents dans le site actif (Phe120 et Phe79 (chaîne B)). Le π -*stacking* est l'interaction attractive non covalente qui peut être formée entre 2 cycles aromatiques.
- Enfin, des interactions électrostatiques entre atomes de charges opposées peuvent être créées entre les acides aminés possédant une chaîne latérale ionisée, et un ligand possédant un groupement ionique.

La liaison hydrogène formée entre le ligand et la cystéine catalytique est la base de l'interaction avec hPRDX5. La force de cette interaction va dépendre de son environnement chimique. Plus particulièrement, elle sera en particulier renforcée si elle est « écrantée » du solvant. En effet comme on peut le voir sur la Figure 73, si le ligand ne s'adapte pas complètement au site et que l'eau pénètre dans cette poche, la force de la liaison H sera diminuée du fait de la compétition avec les molécules d'eau. Au contraire, si le ligand s'adapte au site et qu'il y a désolvatation à l'interface PL, la force de la liaison hydrogène sera renforcée car l'eau ne pourra pas (ou moins) diffuser à l'interface

et entrera en compétition. Dans notre cas, le site de hPRDX5 est majoritairement hydrophobe (Leu, Ile, Phe, Pro, Thr, Gly), ce qui devrait favoriser la désolvatation.

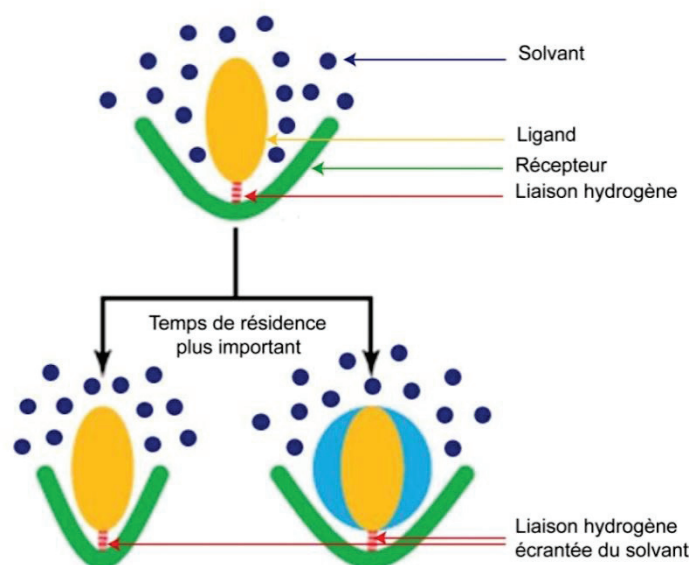
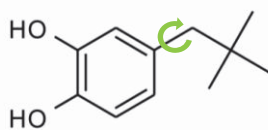


Figure 73. Représentation schématique du principe de liaison hydrogène écrantée du solvant.

L'écrantage du solvant permet de renforcer la force de la liaison H, impliquant des temps de résidence plus importants des complexes PL. L'écrantage est rendu possible par une optimisation du ligand à la taille et la forme du site actif. Figure reproduite et modifiée à partir de Pan et al. *Drug Discovery Today*.

Sur une base intuitive, au vu de la dynamique globale du site actif en interaction avec le ligand, nous avons cherché à rendre la structure du ligand moins rigide en introduisant un degré de liberté dans la structure du ligand. Le but était de potentiellement améliorer l'interaction hydrophobe avec l'hélice Leu116 - Phe120. Nous avons aussi vu précédemment que l'ajout d'un degré de liberté, en permettant une plus grande flexibilité du ligand, pouvait mener à une plus grande dispersion des conformations de l'état lié. Mais les nouvelles conformations rendues possibles par ce degré de liberté ne pourraient-elles pas aussi amener une sélectivité accrue de l'interaction ? Afin d'obtenir des éléments de réponse à cette question, nous avons étudié l'interaction du 4-néopentylcatéchol (4NPC) avec hPRDX5. Partant du 4TBC étudié précédemment, nous avons introduit un degré de liberté qui joue sur le groupement *tert*-butyl en ajoutant un simple $-CH_2$.



4-néo-pentylcatéchol

La nouvelle torsion autorisée par le groupe méthylène additionnel pourrait améliorer son adaptation aux mouvements dynamiques de la protéine. Cela devrait aussi lui permettre d'interagir plus étroitement avec les acides aminés apolaires de l'hélice Leu116 - Phe120 tout en étant un peu plus long. En permettant au ligand d'occuper une place plus importante dans le site catalytique et en lui donnant plus de flexibilité au niveau de la chaîne alkyle, nous cherchons à favoriser la désolvatation du site actif pour renforcer l'écrantage du solvant. Nous attendions donc de cette modification une augmentation du temps de résidence du ligand dans le site actif de hPRDX5.

II. 2. 2 Etude dynamique et statistique de l'interaction 4NPC-hPRDX5

- Simulation de l'interaction du 4NPC par FM

La simulation du 4NPC avec hPRDX5 a été originalement lancée sur une durée de 500 ns, mais la longueur des événements de liaison nous a conduit à pousser cette dynamique jusqu'à atteindre 1 μ s de simulation pour garantir la convergence et renforcer encore l'échantillonnage des différentes conformations.

Au cours de cette simulation, 6 événements de liaison ont été observés. Durant ces périodes de liaison, on peut voir sur le tracé de l'évolution des CV (Figure 74. A) que la valeur de l'angle SG-O_A-O_B (CV3) varie de 60° à 180°. Il y a donc alternance des 2 conformations de liaisons observées au cours de l'étude des dérivés précédents du catéchol (voir I.2 p.108).¹⁸³ Cette observation est confirmée par le tracé des liaisons hydrogènes formées entre SG Cys47 et 4NPC (Figure 74. B), ainsi que par la surface d'énergie libre (FES) du complexe PL tracée en fonction des variables CV1 (distance sur l'axe Z) et CV3 (angle SG-O_A-O_B) (Figure 75). A partir de 800 ns de simulation, la moyenne pondérée de l'énergie libre standard de liaison se stabilise, et tend à la convergence vers une valeur moyenne de $\Delta G_b^\circ = -6.1 \pm 0.8$ kcal.mol⁻¹ soit un $K_D = 60 \pm 60$ μ M. La fluctuation importante de cette énergie vient de la différence d'énergie entre l'état libre et l'état lié. Plus cette différence est importante et plus grande est la fluctuation de l'énergie libre de liaison.

Le *clustering* géométrique a permis la mise en évidence de 4 conformations liées différentes correspondant aux 3 bassins d'interaction visibles sur la FES construite en fonction de CV1 et CV3 (Figure 75). Les deux conformations majoritaires (plus de 60%) sont les conformations correspondant au ligand lié par une liaison H au thiolate de Cys47 (clusters A et B). Elles sont représentatives du bassin d'interaction A+B possédant un angle SG-O_A-O_B de 180°, et différentes de par l'orientation du groupement *néo*-pentyl qui pointe soit vers le solvant (cluster B) soit vers la surface de la protéine (cluster A).

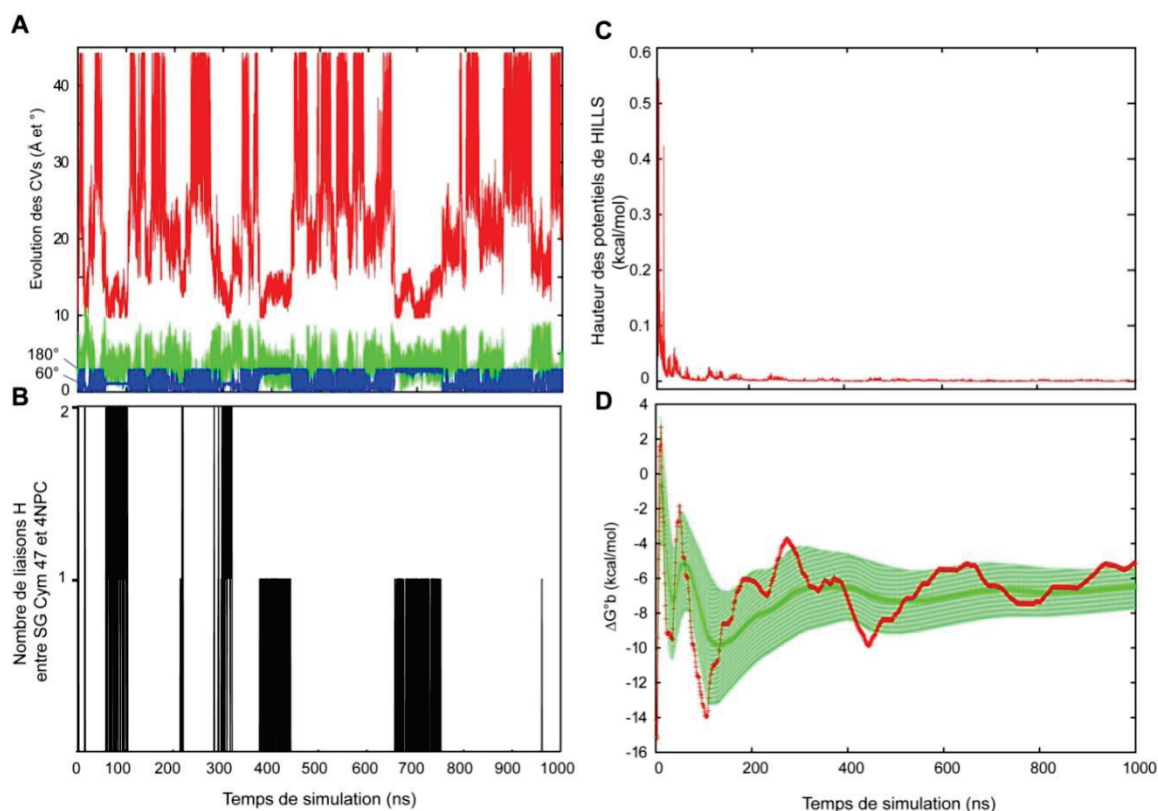


Figure 74. Analyse de la simulation de l'interaction du 4-*néo*-pentylcatéchol avec hPRDX5 par FM.

A) Evolution des CV au cours de la simulation. En rouge (CV1) la distance du ligand sur l'axe Z (Å), en vert (CV2) la distance par rapport à ce même axe (Å), et en bleu (CV3) l'angle SG-O_A-O_B formé par l'atome de soufre de Cys47 et les deux atomes d'oxygène du ligand (°). B) Liaisons hydrogènes formées entre 4NPC et SG Cys47 au cours du temps de simulation. C) Evolution du dépôt des potentiels de HILLS au cours du temps de simulation. D) Evolution de l'énergie libre de liaison du complexe PL au cours de la simulation (rouge), et de l'énergie libre de liaison moyenne pondérée au cours du temps (vert) avec sa fluctuation.

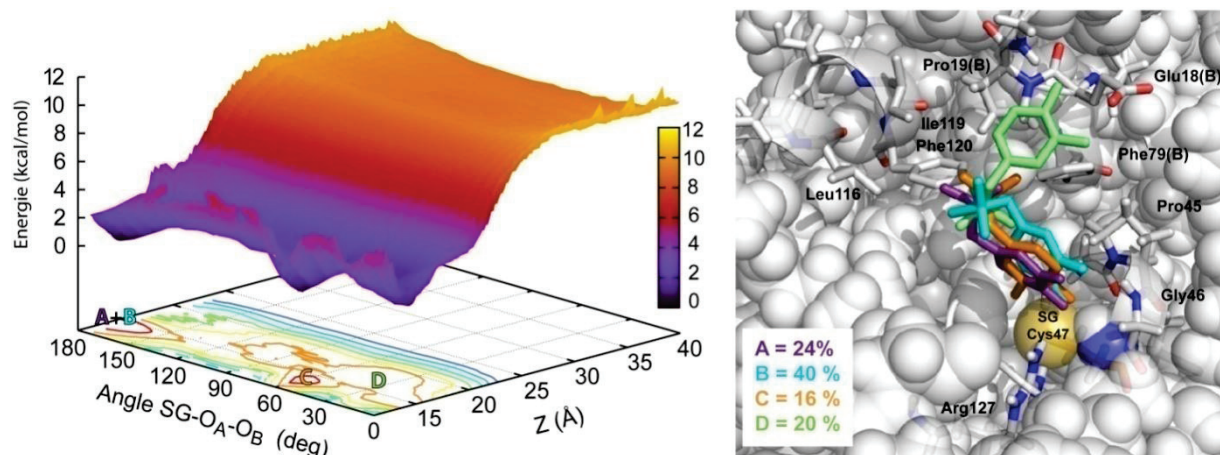


Figure 75. Analyse des 4 modes de liaison du 4-*néo*-pentylcatéchol avec hPRDX5 par FM.

Gauche : FES représentant l'énergie libre du complexe en fonction de CV1 et CV3. Courbes de niveau dessinées avec un intervalle de 1 kcal.mol⁻¹. Droite : Représentation des 4 conformations de liaison obtenues après regroupement géométrique de la trajectoire en utilisant une distance seuil de 2 Å, et représentatives des 3 bassins énergétiques. Cys47 est représentée par une sphère de Van der Waals jaune, et les conformations de liaisons sont colorées en accord avec leur appartenance aux différents bassins d'interactions (A : bleu ; B : violet ; C : orange ; D : vert).

La conformation qui est statistiquement majoritaire est celle présentant le groupement *néo*-pentyl orienté vers le solvant (40%), et le motif catéchol orienté vers l'intérieur du site actif. La conformation liée à SG Cys47 par deux liaisons H (cluster C) n'est pas majoritaire (16%). Elle correspond à la conformation de liaison la moins observée au cours de cette simulation, tandis que la conformation qui décrit le ligand en interaction avec Glu18 (chaîne B) du second monomère, représentent les 20% restants (cluster D). Il y a donc une large prédominance du complexe 4NPC-hPRDX5 lié par le biais d'une seule liaison H par rapport au complexe lié par deux liaisons H, malgré une énergie de même valeur. L'interaction secondaire avec Glu18 (chaîne B) forme un complexe d'énergie plus faible de 1 kcal.mol⁻¹.

- Etude dynamique du site actif

La structure du 4NPC a été proposée dans le but d'améliorer les interactions hydrophobes, mais aussi d'optimiser la flexibilité du ligand à l'intérieur du site actif afin de s'adapter à la dynamique intrinsèque de hPRDX5. Nous nous sommes donc intéressés à la dynamique du site actif au cours des événements d'association et de dissociation du 4NPC. Nous avons regardé l'effet du 4NPC sur la géométrie du site actif durant la simulation FM, et comparé les résultats avec ceux obtenus dans le cas des ligands CAT, 4MEC et 4TBC précédemment étudiés. Pour cela nous avons mesuré au cours du temps de simulation différentes distances interatomiques. Les atomes ont été choisis de manière à ce que les distances puissent refléter l'expansion ou la contraction du site actif. Ainsi, la cystéine catalytique sert de point central pour la mesure de ces 4 distances : SG Cys47 – CB Leu116, SG Cys47 – CB Ile119, SG Cys47 – CZ Phe120 et SG Cys47 – CZ Phe79 (chaîne B).

Afin de relier les variations de distances aux événements de liaison et de dissociation du complexe, la distance entre Cys47 et l'un des oxygènes du ligand (O_A) a aussi été représentée (Figure 76). Si dans le cas du CAT on observe une contraction globale du site actif lors des périodes de liaisons visible par la diminution des distances interatomiques entre Cys47 et les résidus Leu116, Ile119, et Phe120 (encadré orange, Figure 76). Dans le cas du 4MEC et du 4TBC au contraire, le site actif s'élargit en présence du ligand. Cet élargissement est particulièrement visible entre 300 ns et 350 ns dans le cas du 4TBC et autour de 420 ns dans le cas du 4MEC (encadrés violets, Figure 76). Dans les deux cas les ligands ne s'adaptent pas complètement à la dynamique interne du site actif puisque leur association et dissociation provoquent des variations dans les distances interatomiques. Si l'on s'intéresse maintenant à l'interaction du 4NPC, les distances interatomiques ne sont pas affectées par la liaison du ligand à l'intérieur du site actif, même lors de longues périodes comme celle visible entre 650 et 750 ns. Le degré de liberté apporté au système entre le 4TBC et le 4NPC permet donc une meilleure adaptation du ligand dans la cavité catalytique.

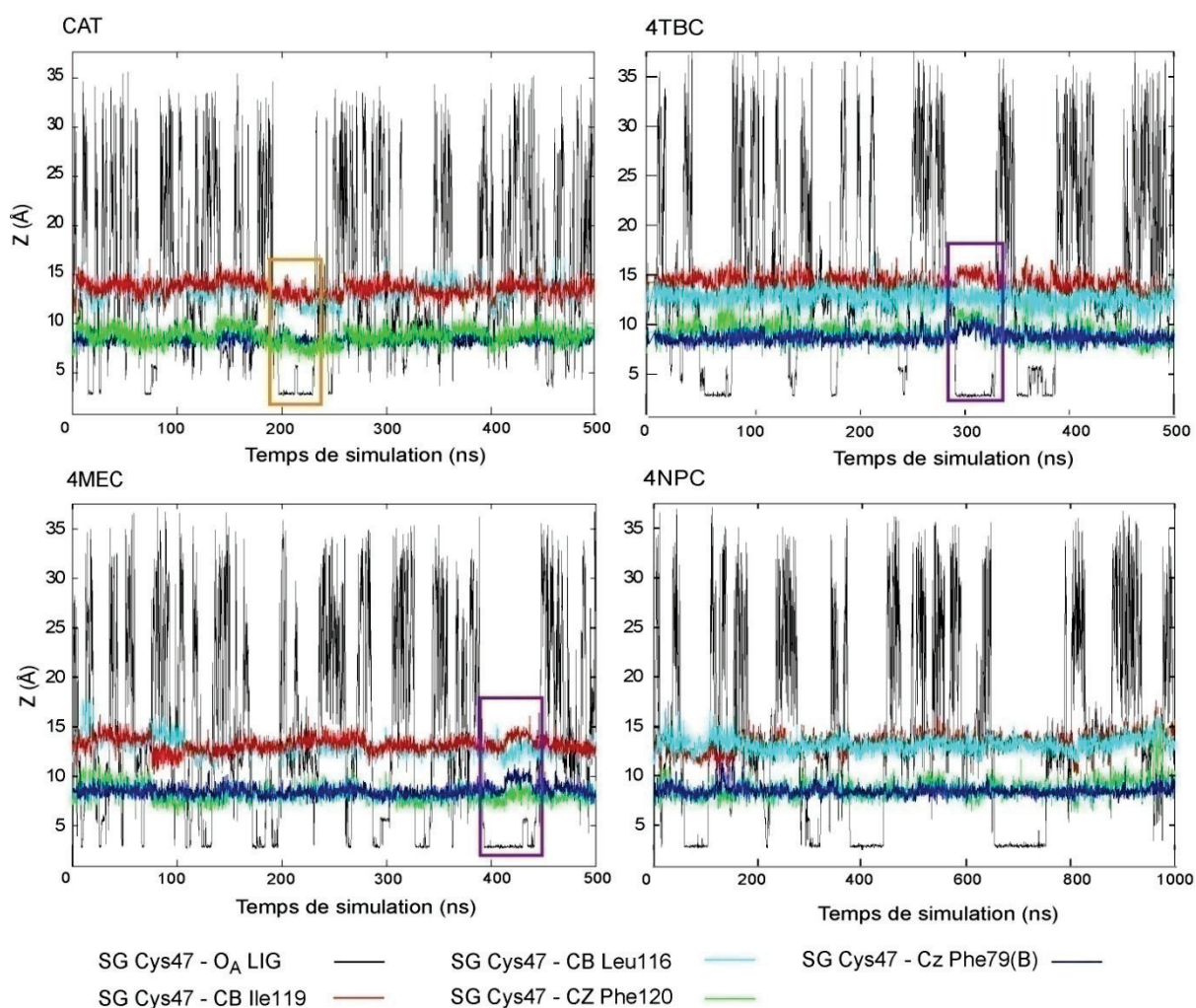
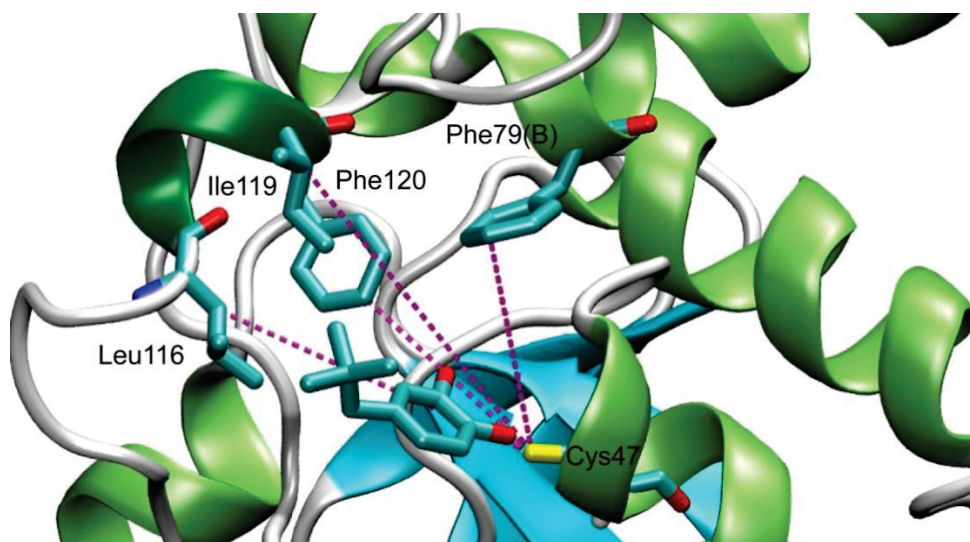


Figure 76. Dynamique interne du site actif par suivi de distances interatomiques en fonction du temps de simulation (ns) pour les 4 ligands suivants : CAT, 4MEC, 4TBC et 4NPC.

La distance entre SG Cys47 et O_A du ligand est tracée en noir. La contraction du site est signalée par un encadré orange, et son expansion par des encadrés violets. Les distances entre SG Cys47 et les acides aminés du site actif considérées, sont tracées avec les couleurs correspondantes signalées.

II. 2. 3 Synthèse du 4-néo-pentylcatéchol

La FM indiquant que le 4NPC aurait une affinité pour hPRDX5 proche du micromolaire, nous avons souhaité obtenir une vérification expérimentale, et vérifier les acides aminés impliqués dans l'interaction. Cette molécule n'étant pas disponible commercialement, nous avons effectué sa synthèse en collaboration avec l'équipe de S. VIDAL à l'Institut de Chimie et Biochimie Moléculaires et Supramoléculaires (ICBMS). La synthèse du 4NPC a été effectuée en 3 étapes, comme décrit en partie III.1. du chapitre II. Le produit désiré a été obtenu sous forme d'un solide incolore avec un rendement global de réaction de 33%. Les intermédiaires réactionnels ainsi que le produit final ont été caractérisés par CCM, RMN, et ESI-MS. Les caractérisations du produit final sont données en Annexe 5.

II. 2. 4 Etude expérimentale par RMN

Nous avons tout d'abord mesuré les CSP induites par le 4NPC sur les signaux de la protéine ^{15}N -hPRDX5 par $^1\text{H}^{15}\text{N}$ -HSQC. Cela a pour but de quantifier la perturbation subie par les acides aminés de hPRDX5 lors de l'interaction, et donc de localiser les acides aminés les plus impactés par l'interaction du 4NPC. D'autre part, en utilisant l'expérience ^1H -STD, nous avons voulu mettre en évidence un ou plusieurs modes de liaison prédominants par la mesure du f_{STD} de chacun des protons aromatiques. Les protons ayant le f_{STD} le plus grand seront ceux qui auront subi le plus gros transfert de saturation et donc seront les plus proches spatialement de hPRDX5.

- Interaction du 4NPC pour hPRDX5 : étude des CSPs par ^{15}N -HSQC

^{15}N -hPRDX5 (110 μM) a été titrée en présence de 4NPC, jusqu'à atteindre une concentration de 1.5 mM de 4NPC, pour un ratio de concentrations [L]/[P] de 14. On note que les CSPs observées sont nettement plus importantes que dans le cas du 4TBC. Les acides aminés ayant les signaux les plus perturbés par l'interaction sont : Ala42, Thr44, Gly46, Cys47, Ser48, Thr50, Leu116, Ile119, Phe120 et Gly148 (Figure 77). On retrouve donc les acides aminés présents dans l'environnement direct de la cystéine catalytique Cys47 (résidus 42 à 50), mais aussi cette fois, les acides aminés constituant l'hélice Leu116 - Phe120 hydrophobe. L'analyse des CSPs pour ces résidus mène à un $\Delta G_b^\circ = -6.6 \pm 0.2 \text{ kcal.mol}^{-1}$ soit un $K_D = 17 \pm 7 \text{ uM}$, du 4NPC pour hPRDX5. Cette valeur est à nouveau en bonne corrélation avec l'affinité prédite par FM. En comparant les CSPs induites dans le cas du 4TBC et du 4NPC, on observe donc que l'amélioration de l'affinité est accompagnée d'une quantité plus importante de perturbation sur l'hélice hydrophobe Leu116 - Phe120, mais aussi au niveau du site actif. Cette augmentation de CSP des résidus situés dans l'environnement de Cys47 peut être le reflet de l'amélioration de la stabilité de la forme liée du complexe PL.

Si l'on considère les agrandissements montrés en Figure 77, on remarque que le système est toujours en échange rapide car aucun dédoublement ou disparition de signal RMN n'est observé.

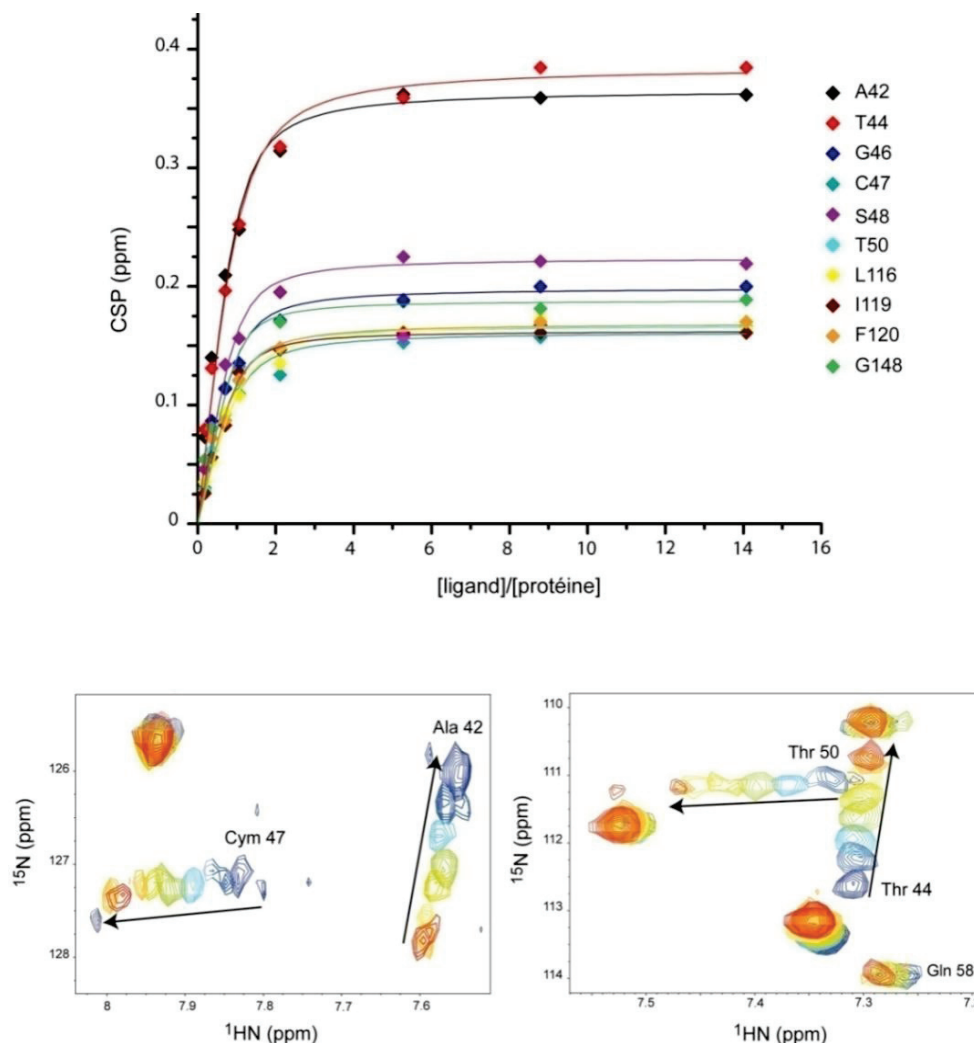


Figure 77. Résultats de l'affinité du 4NPC pour hPRD5 obtenus expérimentalement par RMN en solution.

A) Isothermes de liaison représentés comme la quantité de perturbation de déplacement chimique (ppm) en fonction de la quantité en ligand exprimée comme le rapport des concentrations en ligand et en protéine. Correspondance des points expérimentaux avec des courbes théoriques de fonction hyperbole. B) Deux agrandissements du spectre ^{15}N -HSQC centrés sur les acides aminés Ala42 et Cys47 (gauche), et Thr44 et Thr50 (droite). Superposition des 8 spectres, du bleu (protéine libre) vers le rouge (protéine saturée en ligand).

- Etude de la cartographie d'épitopes par STD

L'expérience STD a été effectuée en présence de 10 μM de protéine, et des ajouts croissants de 4NPC jusqu'à atteindre 1.1 mM. Le calcul des f_{STD} puis le tracé des isothermes de liaison de chaque proton aromatique du ligand en fonction de sa concentration ajoutée nous donne les résultats visibles en Figure 78. Les protons H_5 et H_6 sont les deux protons qui subissent un transfert de

saturation plus important, à hauteur de 90% et 100%, contre seulement 30% pour le proton H₃. Si l'on corrèle ces résultats avec les différents clusters déterminés par FM, on retrouve la prédominance des clusters A et B (Figure 75).

En effet dans ces deux conformations A (bleu) et B (violet), les protons H₅ et H₆ sont orientés vers hPRDX5, proches de Cys47, tandis que le proton H₃ est orienté vers le solvant. Il est donc logique que ce proton subisse un transfert de saturation moins important. Par rapport aux résultats obtenus pour l'interaction du 4TBC avec hPRDX5, on remarque que l'implication de H₃ est nettement plus discriminée, qui pourrait être due à une dispersion moins grande des conformations liées que dans le cas du 4TBC.

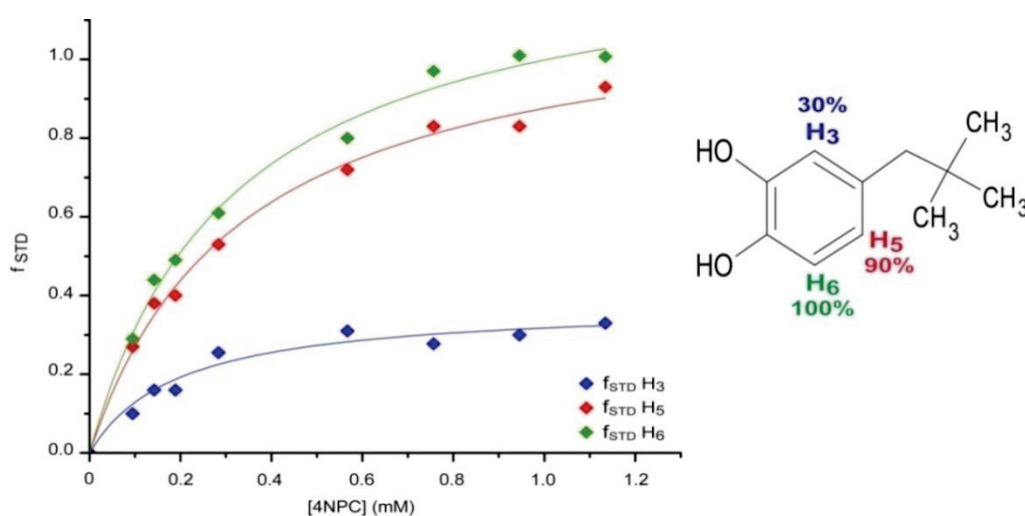


Figure 78. Résultats obtenus par expériences RMN STD pour l'interaction de 4NPC avec hPRDX5.

A) Courbes représentant f_{STD} pour les protons aromatiques de 4NPC en fonction de la concentration en ligand. Correspondance des points expérimentaux avec une courbe théorique (traits pleins) calculée par une équation de type hyperbole $y = A x / B + x$. B) Superposition des spectres 1D et STD au niveau des protons aromatiques pour une concentration en 4NPC de 1.2 mM.

II. 3. Optimisation du 4-néo-pentylcatéchol par MD

Les résultats obtenus dans le cas du 4NPC sont encourageants sur le caractère prédictif des protocoles de MD, mais l'affinité de ce ligand pourrait-elle être encore améliorée ? Dans la conception du 4NPC nous avons réussi à améliorer les interactions de vdW entre le ligand et l'hélice Leu116 - Phe120, induisant une meilleure stabilité du complexe lié et menant à une affinité de l'ordre de 20 μ M. L'ajout du degré de liberté sur la chaîne aliphatique en position 4 du cycle aromatique a permis au fragment 4NPC de s'adapter à la dynamique intrinsèque du site actif de hPRDX5.

Nous avons vu en chapitre I que pour développer un *hit*, il est possible d'ajouter des groupements fonctionnels afin de chercher à améliorer les propriétés ADMET ou la LE en augmentant le nombre d'interactions entre le ligand et la protéine. Le but est de stabiliser la forme liée et d'améliorer le temps de résidence moyen \bar{t}_R du ligand. Nous avons mis en évidence une interaction secondaire récurrente avec Glu18 de la chaîne B de l'homodimère. Nous avons alors cherché à développer le fragment 4NPC afin qu'il puisse se lier à la fois à Cys47 et à Glu18 (chaîne B), tout en conservant les interactions de vdW formées avec l'hélice Leu116 - Phe120. L'augmentation de la taille du ligand pourrait aussi permettre de discriminer l'un des 2 modes de liaison par rapport à l'autre. Plusieurs étapes de développement ont été effectuées afin d'observer l'influence de chacune des modifications apportées à la structure du ligand.

II. 3. 1 Etude du 6-(néopentyl)biphenyl-3,4-diol et de dérivés

- Etude de l'affinité du 6-(néopentyl)biphenyl-3,4-diol par FM

Nous avons vu précédemment que dans la cavité du site actif se trouvaient des acides aminés hydrophobes, dont des acides aminés aromatiques comme Phe120 et Phe79 (chaîne B). Pour cette raison, nous avons choisi d'introduire un second cycle benzénique afin de former des interactions attractives non covalentes de type vdW avec ces résidus grâce aux électrons π délocalisés des cycles aromatiques (le π -stacking, Figure 79). De plus, lors du criblage RMN effectué en 2010 au laboratoire, plusieurs molécules de type biphenyl avaient été trouvées comme ayant une possible affinité pour hPRDX5.⁹¹

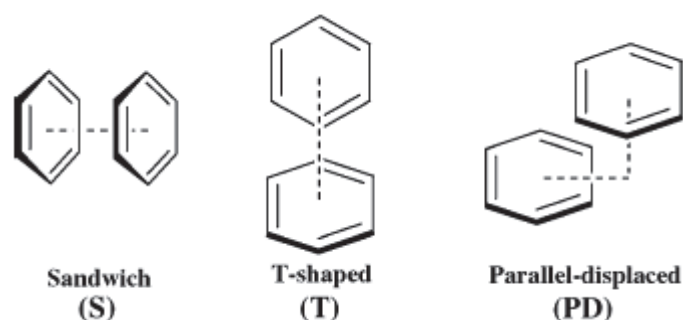
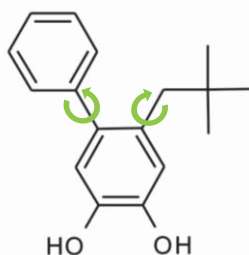


Figure 79. Différentes interactions non covalentes attractives possibles entre deux cycles aromatiques : le « π -stacking ».

π -stacking entre deux cycles benzéniques, selon leurs position l'un par rapport à l'autre : l'interaction sandwich (S), l'interaction en T (T), et l'interaction parallèle décalée (PD). L'interaction S semble être la plus favorable car le recouvrement est maximal, mais elle est la moins observée expérimentalement. Une étude théorique de l'énergie d'interaction de ces trois configurations de dimère benzénique indiquerait que les configurations S et DP seraient plus favorables énergétiquement ($-2.6 \text{ kcal.mol}^{-1}$) que la configuration S ($-1.7 \text{ kcal.mol}^{-1}$).¹⁸⁹

L'ajout de ce cycle aromatique peut mener à une diminution de la LE en comparaison du 4NPC si ΔG_b° n'est pas améliorée, puisque 6 atomes lourds sont ajoutés. Cependant, cela pourrait être compensé par un gain plus important d'affinité grâce aux interactions de vdW aromatiques possibles en plus de l'interaction hydrophobe existante entre l'hélice Leu116 - Phe120 et le groupement *néo*-pentyl en position 4 du catéchol. L'ajout de nouvelles interactions pourrait conduire à une plus grande stabilisation du complexe PL ainsi qu'une augmentation du temps de résidence pour cette structure qui s'éloigne du type « fragment » et qui pourrait former plus d'interactions de surface.



6-(*néo*-pentyl)biphényl-3,4-diol

L'analyse de la simulation FM du 6-(néopentyl)biphényl-3,4-diol (6NBD) avec hPRDX5 nous indique que ce ligand possède une structure tridimensionnelle trop grande pour le site actif du fait de l'arrangement des deux noyaux aromatiques plans.

Un seul évènement de liaison a pu être enregistré, d'une durée de 100 ns au cours duquel un échange de conformation a eu lieu. Nous avons de plus remarqué que lors de la liaison à Cys47 par une liaison H (angle SG-O_A-O_B de 180°), le passage brusque à la seconde conformation de liaison par création de la seconde liaison H (angle SG-O_A-O_B de 60°) se faisait de manière instantanée.

Le ΔG_b° de ce complexe s'assimile à l'affinité du dérivé 4NPC étudié précédemment, voir légèrement meilleure : $\Delta G_b^\circ = -6.9 \pm 0.4 \text{ kcal.mol}^{-1}$ soit $K_D = 14 \pm 10 \mu\text{M}$.

La FES du complexe du 6NBD avec hPRDX5 en fonction des CV 1 et CV 3 montre une différenciation en termes d'angle (bassins entre 30 et 60° puis autour de 120°), mais pas en termes de distance sur l'axe Z du fait de la définition du centre de masse du 6NBD (Figure 80). En effet, le centre de masse de 6NBD se trouve à une distance d'environ 17 Å sur Z s'il est en interaction avec Cys47, de même que s'il est en interaction avec Glu18 (chaîne B). L'analyse de la trajectoire n'a montré aucune preuve d'interaction entre le ligand et les résidus aromatiques Phe120 et Phe79 (chaîne B). Notre hypothèse est donc que le ligand serait trop rigide pour se positionner dans le site actif de façon à interagir à la fois avec l'hélice Leu116 - Phe120 par le biais du groupement *néo*-pentyl et avec les résidus aromatiques. Une fois lié à Cys47, sa taille et sa rigidité lui empêchent de modifier son orientation de façon à optimiser les interactions formées ou à en former de nouvelles.

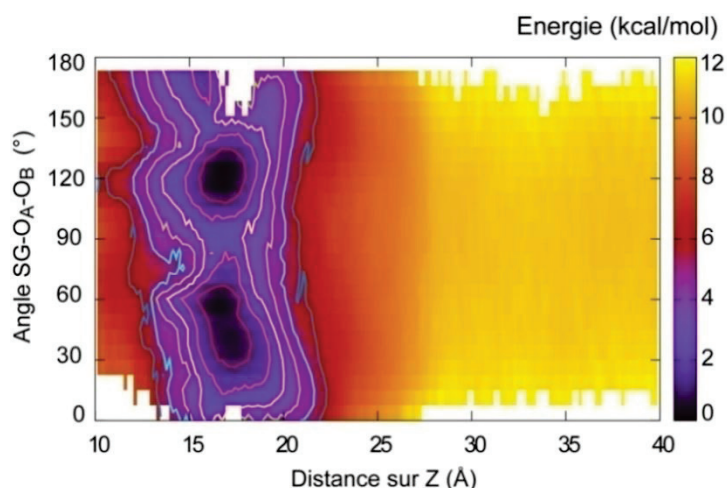
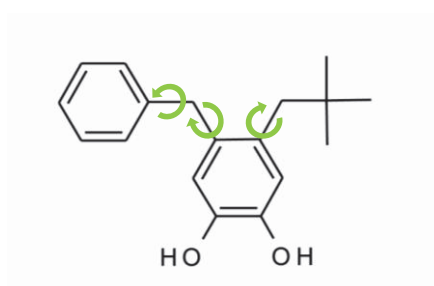


Figure 80. Surface d'énergie libre du complexe 6NBD-hPRDX5 au cours de la simulation FM.

FES tracée en fonction des variables collectives CV1 (Z (Å)) et CV3 (angle $SG-O_A-O_B$ formé par l'atome de soufre de Cys47, et les atomes d'oxygène O_A et O_B de 6NBD). Courbes de niveau dessinées en traits pleins avec un pas de 1 kcal.mol^{-1} .

- Etude du 4-benzyl-5-néopentylcatéchol par FM

Afin de diminuer la rigidité du ligand par rapport au 6NBD, nous avons alors choisi d'introduire un groupement méthylène afin de rendre possible la rotation du second cycle benzénique. Ayant eu un impact positif dans le cas du ligand 4NPC par rapport au ligand 4TBC, nous avons réitéré l'expérience sur ce ligand 4-benzyl-5-néopentylcatechol (4B5NPC).



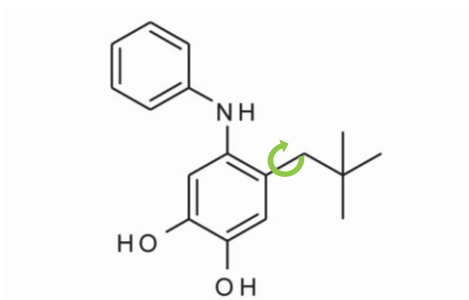
4-benzyl-5-*néo*-pentylcatéchol

Dans le cas présent, après 600 ns de simulation 4 évènements de liaison ont été détectés : 3 courts évènements ($\leq 10 \text{ ns}$) avec formation d'une seule liaison H, et une période de 50 ns pendant laquelle le ligand était lié par deux liaisons H. Nous avons aussi observé lors de l'analyse de la trajectoire que le ligand 4B5NPC formait de nombreuses interactions faibles en surface de hPRDX5. En calculant l'énergie libre de liaison de ce complexe pour hPRDX5, on obtient une énergie libre de liaison du même ordre de grandeur, mais plus faible que celle obtenue pour le ligand 6NBD : $\Delta G_b^\circ = -6.3 \pm 0.8 \text{ kcal.mol}^{-1}$ soit un $K_D = 50 \pm 50 \mu\text{M}$.

L'introduction de ce degré de liberté n'a donc pas eu pour effet de faciliter l'adaptation du ligand à l'intérieur de la cavité. La nouvelle torsion introduite induit une augmentation significative de la taille du ligand qui a encore plus de mal à se positionner dans le site actif, et amène de nombreuses interactions de surface sans effet bénéfique. De plus, le noyau aromatique n'a montré la formation d'aucune interaction de type aromatique (π -stacking) avec ce degré de liberté supplémentaire.

- Etude du 4-néopentyl-5-(phénylamino)catéchol

Après avoir introduit un degré de liberté supplémentaire entre les noyaux benzéniques, sans effet bénéfique sur l'affinité PL, nous avons modifié la nature de l'atome autorisant ce nouveau degré de liberté. Nous avons introduit un atome d'azote qui permettrait d'augmenter la lipophilie du ligand en comparaison du dérivé précédent, et d'apporter de nouvelles propriétés électroniques dues à la réactivité de l'azote.



4-néo-pentyl-5-(phénylamino)catéchol

1 400 ns (1.4 μ s) de simulation ont été nécessaires à l'observation de 5 évènements de liaison dont 3 d'environ 120 ns entre le 4-néopentyl-5-(phénylamino)catéchol (4N5PC) et hPRDX5, indiquant une augmentation significative du $\bar{\tau}_R$ apparent. La longueur de ces évènements de liaison a aussi pour conséquence un très bon échantillonnage de l'état lié et donne lieu à une énergie libre de liaison moyenne à la convergence de $\Delta G_b^\circ = - 8.0 \pm 1.0$ kcal.mol⁻¹ soit un $K_D = 3 \pm 1$ μ M.

La surface d'énergie libre du complexe en fonction des variables collectives CV1 et CV3 montre une énergie bien différenciée des 2 minima énergétiques : une énergie plus faible de 2 kcal.mol⁻¹ au niveau de la liaison avec la Cys47 en comparaison avec l'interaction avec Glu18 (chaîne B) (Figure 81). Considérant les différentes conformations de liaison mises en évidence par le clustering géométrique, on peut voir que le noyau aromatique ne forme toujours aucune interaction avec d'autres acides aminés aromatiques (Figure 81). Cependant, le remplacement du groupement méthylène par un groupement amine a permis une amélioration de l'affinité pour hPRDX5 de l'ordre de 1 kcal.mol⁻¹ par rapport au ligand 6NBD étudié précédemment.

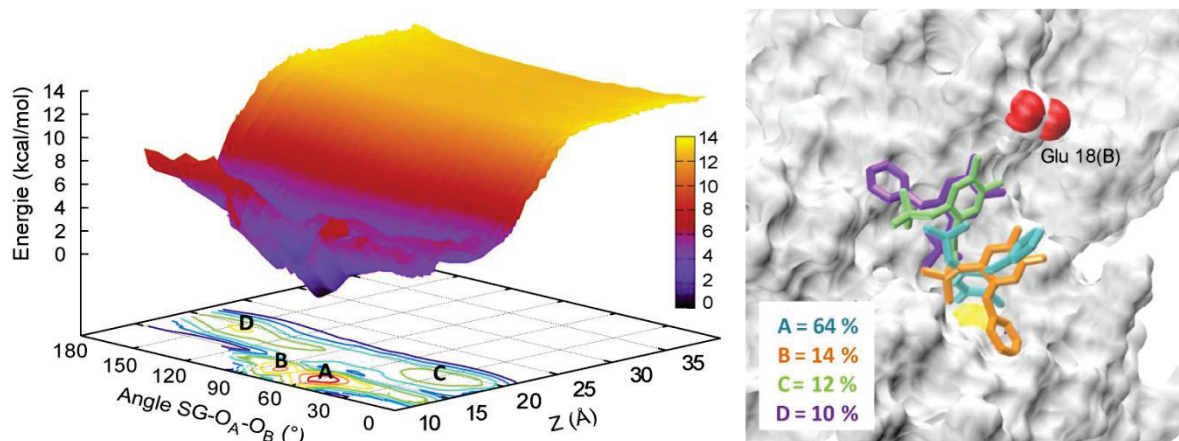


Figure 81. Analyse de l'interaction du 4-néopentyl-5-(phénylamino)catéchol avec hPRDX5.

Gauche : FES représentée en fonction de la distance sur l'axe Z (Å) (CV1) et de l'angle SG-O_A-O_B (°) (CV2). Courbes de niveaux dessinées avec un pas de 1 kcal.mol⁻¹. Droite : les 4 conformations de liaisons représentatives des minima énergétiques avec leur proportion relative de présence pour cette simulation. Atome SG Cys47 représenté sous forme d'une sphère de van der Waals jaune.

Tout comme nous l'avons fait dans le cas du 4NPC, nous avons étudié l'impact de l'interaction de 4N5PC sur la dynamique du site actif afin de voir si l'ajout du cycle benzénique n'induit pas une déformation du site actif. Il en résulte que le ligand s'adapte tout autant que le 4NPC malgré une structure tridimensionnelle plus volumineuse. En effet les distances entre les acides aminés entourant le site actif restent stables durant les périodes de liaison (Figure 82). Aucun nouveau type d'interaction n'a été détecté lors de cette simulation.

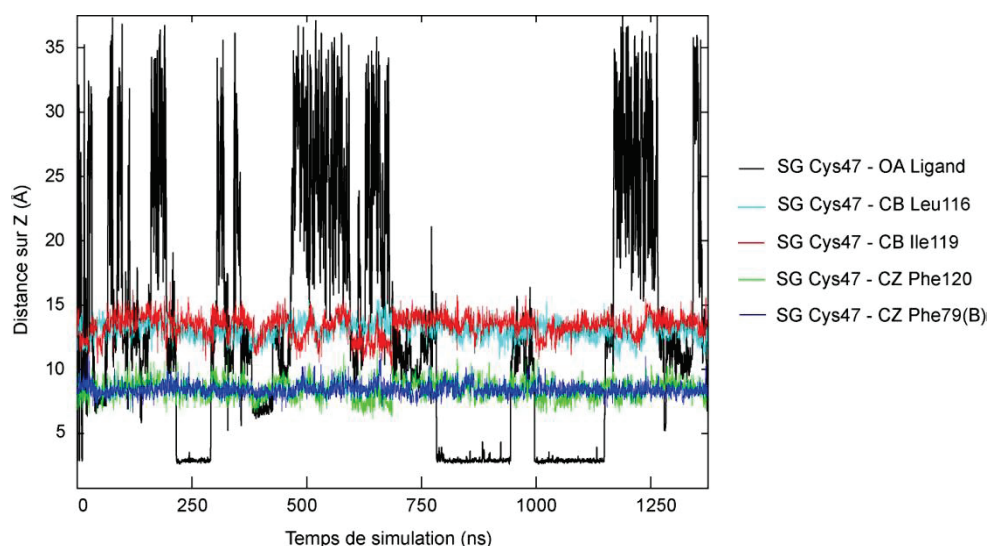
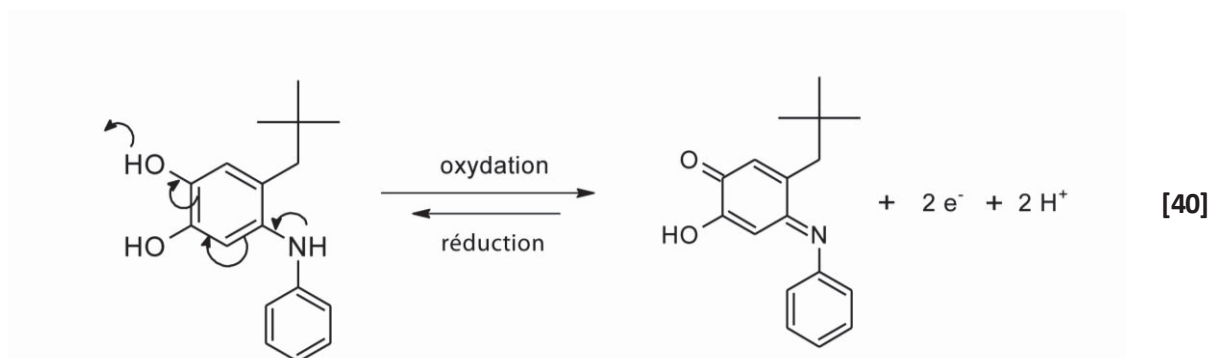


Figure 82. Dynamique du site actif en présence du 4-néopentyl-5-(phénylamino)catechol.

Distances mesurées entre Cys47 et les acides aminés Leu116, Ile119, Phe120 et Phe79 (chaîne B) (Å) et colorées comme indiquées sur le graphique. Distance entre la cystéine catalytique et le ligand selon l'axe Z de l'entonnoir en ligne noire, et mise en évidence des événements de liaison.

L'augmentation de l'affinité observée est donc peut être uniquement due à l'augmentation de la taille du ligand, et aussi peut être du fait de l'écrantage du solvant. Cette amélioration peut aussi venir de la formation privilégiée de la conformation liée par le biais de deux liaisons H (conformation A, Figure 81) en comparaison des autres dérivés qui étaient jusqu'à présent en échange entre deux conformations lors de leur liaison avec Cys47.

Au vu de la forte affinité estimée par FM de ce ligand pour hPRDX5, nous avons souhaité étudier son interaction par RMN afin de quantifier les CSPs induites. Cette molécule n'étant pas disponible commercialement, elle a été synthétisée à l'ICBMS de Lyon, dans le cadre de notre collaboration avec le Dr S. Vidal (Figure 40 p.90). Le produit a été isolé sous forme d'un mélange des formes oxydée et réduite, sous forme de solide rouge brique. Cette propriété rédox n'avait pas été vraiment prévue à l'origine. L'atome d'azote favorise la délocalisation des électrons du cycle catéchol et va induire la formation de la forme oxydée quinoneimine selon l'équilibre **40** d'oxydoréduction ci-après.



Dans notre protocole RMN, l'enzyme hPRDX5 doit être sous forme réduite. Le TCEP maintient réduit le 4N5PC. Cependant, même en présence d'un excès de réducteur, l'oxygène de l'air suffit à consommer l'excès de réducteur et les échantillons RMN doivent être préparés sous atmosphère inerte (N₂ ou Argon). L'oxydation du ligand par l'oxygène pouvant mener à l'oxydation partielle de la protéine, les quantités de ligand et de protéine réduites en solution ne peuvent être considérées cette fois comme constantes.

De plus, des précipités ont été observés contenant du 4N5PC sous forme oxydée. L'affinité du 4N5PC pour hPRDX5 calculée par ¹H-¹⁵N HSQC comme étant de $K_D = 100 \pm 30 \mu\text{M}$ est donc, selon notre hypothèse, sous-estimée puisque la quantité réelle en ligand est bien inférieure à celle introduite en solution. Cela laisse à penser que l'affinité réelle de ce ligand pour hPRDX5 peut atteindre le micromolaire.

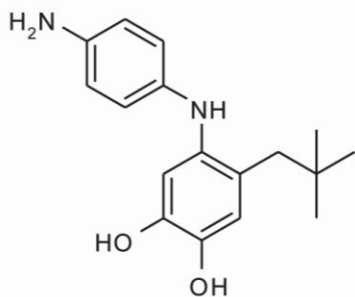
II. 3. 2 Etude de dérivés du 4-néopentyl-5-(phénylamino)catechol par FM

Le dérivé 4N5PC ayant donné lieu par FM à une affinité estimée améliorée pour hPRDX5 par rapport au ligand 4NPC tout en permettant de favoriser un mode de liaison particulier, nous avons décidé de conserver la structure de ce ligand pour de futures optimisations. Pour aller plus loin, nous avons introduit de nouvelles fonctions chimiques en position para du cycle benzénique afin de diminuer son hydrophobicité. Ont été étudiés une fonction amine primaire R-NH₂, une fonction carboxylate R-COO⁻, et une fonction méthylammonium R-CH₂-NH₃⁺. Ces molécules ont été modélisées et leur affinité pour hPRDX5 estimée par FM.

- Etude des ligands 4-[(4-aminophényl)amino]-5-(néo-pentyl)catéchol et 4-[[2-(néo-pentyl)-4,5-dihydroxyphényl]amino]benzoate par FM

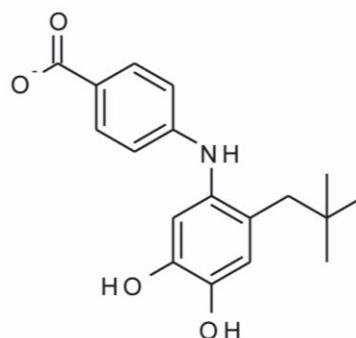
Nous avons tout d'abord introduit un groupement amine en position para de l'aminophényl. Il faut noter que cette amine étant rattachée au cycle benzénique, elle possède les propriétés de l'aniline. Son doublet non liant étant alors délocalisé, à pH=7.4 le groupement amine est sous forme basique non protonée NH₂ (pKa = 4.6). L'aniline étant une base faible à la fois donneuse et accepteuse de liaisons H, elle pourrait hypothétiquement former des interactions ioniques avec les acides aminés hydrophiles acides (Asp, Glu, Cys) ou basiques (Arg, His, Lys), ainsi que des liaisons H. Nous avons par la suite introduit à cette même position un groupement carboxylate car à pH = 7.4 un acide carboxylique est ionisé et présent sous forme COO⁻. Le groupement carboxylate étant une base faible accepteuse de liaisons H, hydrophile et électriquement chargé, il pourrait former des liaisons H, ou des interactions de type coulombiennes avec des résidus Arg.

Les ligands que nous avons étudiés sont le 4-[(4-aminophényl)amino]-5-(néo-pentyl)catéchol (Hit 1) et le 4-[[2-(néo-pentyl)-4,5-dihydroxyphényl]amino]benzoate (Hit 2). L'interaction de ces ligands avec hPRDX5 ont été simulées par FM, et les résultats d'affinité ont été regroupés dans le Tableau 9 ci-dessous.



4-[(4-aminophényl)amino]-5-(néo-pentyl)catéchol

Hit 1



4-[[2-(néo-pentyl)-4,5-dihydroxyphényl]amino]benzoate

Hit 2

Tableau 9. Résultats d'affinité des ligands Hit1 et Hit2 pour hPRDX5 estimés par FM

Ligand	Durée de la simulation (ns)	ΔG°_b (kcal/mol)	K_D (μM)
Hit 1	500	- 5.7 \pm 0.5	100 \pm 150
Hit 2	500	- 4.8 \pm 0.2	330 \pm 150

Dans le cas du Hit 1, seulement 3 évènements de liaison (entre 10 et 50 ns) ont été observés, avec un \bar{t}_R plus faible que celui observé pour le ligand sans amine. Pour le Hit 2, 5 périodes de liaison ont eu lieu, durant lesquelles les deux conformations de liaison sont visibles (à 1 et 2 liaisons hydrogènes). Les évènements sont de courte durée, (10 à 50 ns), tout comme le Hit 1, et l'affinité pour hPRDX5 est moins bonne. En cartographiant l'énergie libre de liaison de ces ligands en interaction avec hPRDX5 en fonction des variables collectives CV1 et CV3, on retrouve dans les deux cas 2 bassins d'énergie distincts correspondant aux interactions avec Cys47 (par le biais de 1 ou 2 liaisons H) et Glu18 (chaîne B). Cependant dans le cas du Hit 1, le bassin d'énergie correspondant à l'interaction secondaire avec Glu18 (chaîne B) est de plus faible énergie, ce qui veut dire que l'interaction de ce ligand avec Glu18 (chaîne B) est plus forte que son interaction avec Cys47 (Figure 83). Le peu de liaisons formées avec Cys47 au cours de la simulation FM n'aura sans doute pas permis un échantillonnage assez important de ce bassin énergétique au vu de l'espace non visité (Figure 83) et des répulsions électrostatiques possibles entre le groupement carboxylate et les résidus Glu et Asp de hPRDX5.

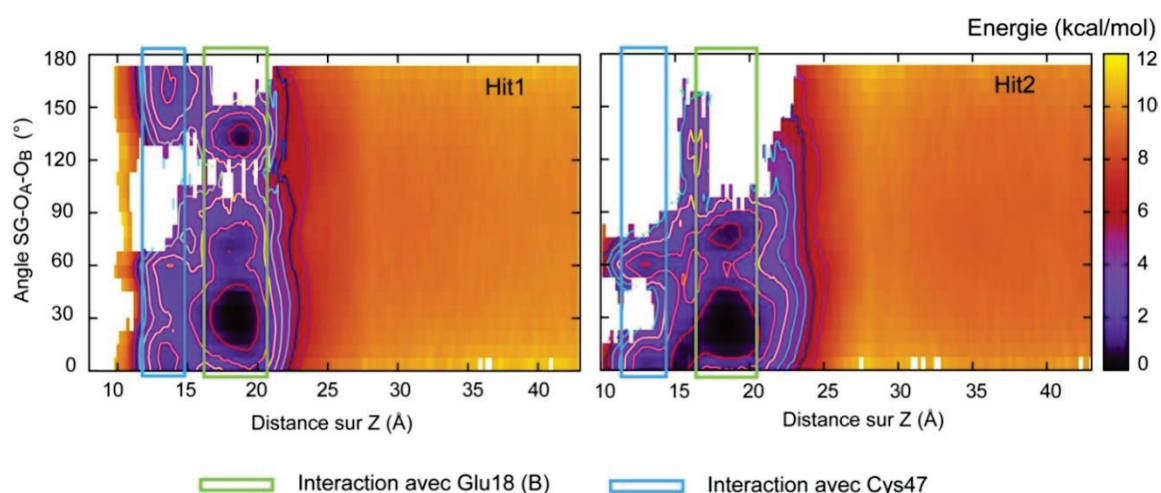


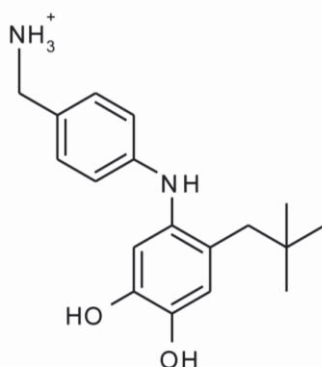
Figure 83. Surfaces d'énergie libre de liaison des ligands Hit1 et Hit2 en interaction avec hPRDX5 lors de la simulation FM.

FES tracées en fonction des variables collectives CV1 (Z (Å)) et CV3 (angle SG-O_A-O_B formé par l'atome de soufre de Cys47, et les deux atomes d'oxygène O_A et O_B du ligand). Courbes de niveau dessinées en traits pleins avec un pas de 1kcal.mol⁻¹. Zones non explorées par le ligand plus importantes dans le cas du Hit 2, probablement causées par des répulsions électrostatiques en surface de hPRDX5.

De même, l'analyse de la FES en fonction des CV1 et CV3 nous montre que les bassins d'énergie correspondants à l'interaction du Hit 2 avec Glu18 (chaîne B) sont de même énergie que ceux correspondant à l'interaction avec Cys47, voire plus faible. L'ajout des groupements $-\text{COO}^-$ et $-\text{NH}_2$ n'apportent pas de nouvelles interactions à l'intérieur du site actif permettant d'améliorer la stabilisation de l'état lié, et tendent même à délocaliser l'interaction vers Glu18 (chaîne B), en surface de hPRDX5. Enfin, la taille de ces ligands de type « *drug-like* » est telle que l'échantillonnage de l'espace est limité, tandis que les interactions de surface prennent plus d'importance (Figure 83). Cela peut expliquer la diminution de l'affinité apparente observée de ces ligands Hit 1 et Hit 2 pour hPRDX5.

- Etude de l'interaction du (4-{[2-(néo-pentyl)-4,5-dihydroxyphényl]amino}phényl)méthanaminium avec hPRDX5 par FM

Nous avons choisi d'introduire ici un groupement chargé positivement afin d'avoir la possibilité de former des interactions ioniques plus spécialement avec la chaîne latérale de Glu18 (chaîne B), tout en maintenant le groupement catéchol dans la cavité catalytique. Notre choix s'est porté sur le groupement méthanaminium pour le pK_a de l'amine primaire ($pK_a = 10.7$), ainsi que pour sa praticité de synthèse. Le ligand étudié est le (4-{[2-(néo-pentyl)-4,5-dihydroxyphényl]amino}phényl) méthanaminium (Hit 3).



(4-{[2-(néo-pentyl)-4,5-dihydroxyphényl]amino}phényl)méthanaminium

Hit 3

La simulation FM de 700 ns du Hit 3 avec hPRDX5 a permis de comptabiliser 6 évènements de liaison, avec un $\bar{\epsilon}_R$ moyen apparent plus faible que dans le cas du 4N5PC (Figure 84.A). L'énergie libre absolue de liaison estimée en fin de simulation pour le complexe Hit3-hPRDX5 est proche de celle obtenue pour le dérivé 4N5PC : $\Delta G_b^\circ = -8.2 \pm 1.0 \text{ kcal.mol}^{-1}$ correspondant à un $K_D = 3 \pm 3 \text{ }\mu\text{M}$. Ce résultat présente une fluctuation de l'énergie libre de liaison moyenne, qui provient d'un évènement de liaison visible entre 500 ns et 600 ns. Nous avons pu remarquer que la fluctuation de l'énergie est

plus grande pour les ligands ayant des évènements de liaison de longue durée. Dans le cas présent, l'énergie moyenne pondérée au cours du temps reste stable en fin de simulation et converge vers cette valeur moyenne (Figure 84.B).

En regardant la cartographie de l'énergie libre de liaison en fonction des variables CV1 et CV3, on peut distinguer de manière plus nette les bassins d'énergie correspondant respectivement à l'interaction avec Cys47 (bassins A et B) et avec Glu18 (chaîne B), (bassin C) maintenant séparés de 4 kcal.mol⁻¹. Si l'on regarde en fonction de l'angle SG-O_A-O_B, les deux conformations de liaison (A et B) qui ont été bien échantillonnées sont de même énergie, ce qui ne laisse pas présager d'une meilleure sélectivité de la conformation de liaison (Figure 84.C). Enfin, on peut remarquer que malgré sa taille, plus grande que le ligand 4NPC par exemple, le Hit 3 a pu échantillonner la quasi-totalité de l'espace lors de la simulation FM. Par *clustering* géométrique, on trouve les 3 conformations de liaison observées pour l'ensemble des autres dérivés catéchols : les deux conformations de liaison A et B avec Cys47 (par le biais de une ou deux liaisons hydrogènes), et la conformation de liaison C en interaction avec Glu18 (chaîne B) (Figure 84.D). La conformation liée par une seule liaison hydrogène (A) est dans le cas présent prédominante en terme de présence puisqu'elle représente 47% des conformations liées, contre 16% pour le second mode de liaison observé avec Cys47 par le biais de 2 liaisons H. Le duster représentatif de l'interaction secondaire avec Glu18 (chaîne B) reste fortement représenté avec 37% des conformations liées. Il est aussi possible de visualiser la conformation du ligand en interaction avec Cys47, tandis que le groupement méthaniminium est en interaction avec Glu18 (chaîne B), ce qui était la conformation recherchée mais qui reste pour le moment en très faible proportion et n'est donc pas visible dans les clusters.

En comparaison du ligand catéchol initialement étudié, les modifications structurales apportées jusqu'à l'étude de ce Hit 3 par FM ont donné lieu à une amélioration de l'affinité pour hPRDX5 d'un facteur 1000 environ, atteignant à présent un K_D de l'ordre du micromolaire, estimé par FM.

Pour aller plus loin, il faudrait mettre en évidence d'autres interactions de surface possibles dans l'environnement du site actif, sans pour autant déstabiliser la forme liée du complexe.

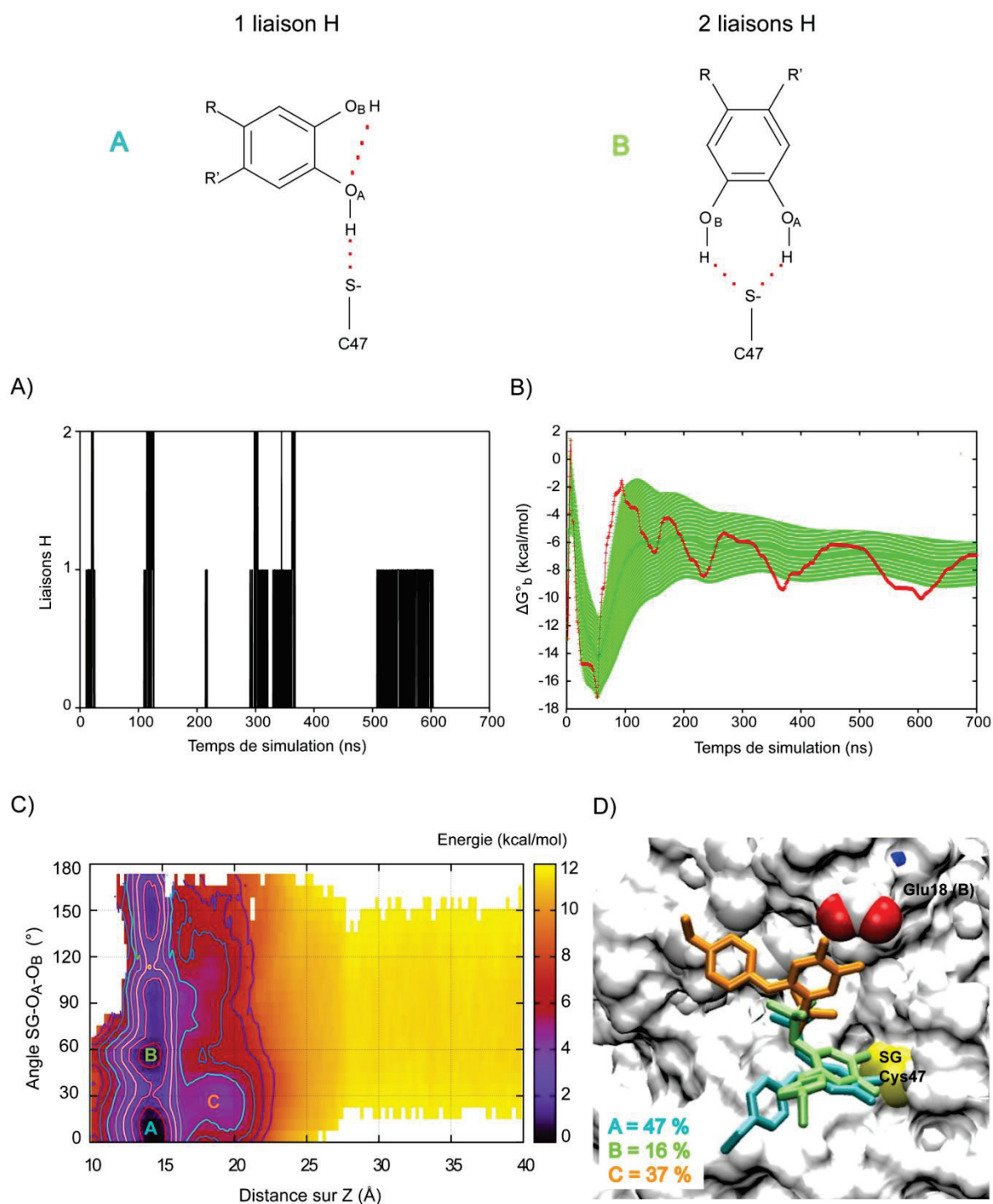


Figure 84. Analyse des modes de liaisons majoritaires du Hit 3 dans le site catalytique de hPRDX5, et visualisation de l'énergie de liaison du complexe associé.

A) Liaisons hydrogènes formées entre Hit 3 et SG de Cys47 au cours du temps de simulation (ns). B) Evolution de l'énergie libre de liaison du complexe PL au cours de la simulation (rouge), et de l'énergie libre de liaison moyenne pondérée au cours du temps (vert) avec sa fluctuation. C) FES du complexe Hit 3 – hPRDX5 tracée en fonction des variables collectives CV1 (Z (Å)) et CV3 (angle SG-O_A-O_B formé par l'atome de soufre de la cystéine catalytique, et les deux atomes d'oxygène O_A et O_B du ligand). Courbes de niveau dessinées en traits pleins avec un pas de 1 kcal.mol⁻¹. D) Les 3 conformations de liaisons représentatives des minima énergétiques visibles sur la FES et leur proportion relative de présence pour cette simulation. Atome SG de Cys47 représenté sous forme d'une sphère de van der Waals jaune. Groupements : R= aminophénylméthaniminum ; R' = neopentyl.

II. 4. La FM pour l'optimisation des petites molécules bioactives

Après avoir démontré dans le premier chapitre de résultats que la FM permettait d'une part l'estimation de l'affinité de manière quantitative, et d'autre part l'étude statistique des modes de liaison et des interactions mises en jeu, l'un des objectifs de ma thèse était d'utiliser ces données pour optimiser les molécules fragments étudiées afin d'en améliorer l'affinité pour hPRDX5. En effet, le nombre de traitements existant contre l'inflammation post-ischémique cérébrale étant limité, il est aujourd'hui capital d'identifier de nouveaux inhibiteurs comme potentiels médicaments.

Nous avons vu dans les résultats présentés dans cette partie que par l'optimisation du fragment catéchol, nous avons pu atteindre une affinité micromolaire pour hPRDX5 grâce aux informations données par la FM (Tableau 10 et Figure 85). A partir de l'étude de fragments simples tels que les cyclohexane-1,2-diols et le gallate d'éthyle, nous avons vu que la présence du cycle aromatique et de la position 1,2 des groupements hydroxyles sur ce cycle était capitale pour maintenir une interaction avec Cys47, afin de conserver une analogie avec l'état de transition de la réaction de réduction de H₂O₂ catalysée par hPRDX5 (voir chapitre I, p.72).

Par la suite, nous avons constaté que l'on obtenait de meilleures affinités lorsque le groupement en position 4 du cycle aromatique était apolaire, typiquement un groupement alkyle. En effet lorsque le ligand est en liaison avec Cys47, ce substituant peut-être orienté face à l'hélice Leu116 – Phe120 délimitant l'un des côtés du site actif, pour former des interactions de vdW de type hydrophobes qui ont pour conséquence de stabiliser l'état lié du complexe. Ces premières expériences ont menées à la synthèse du 4NPC, pour lequel l'introduction d'un degré de liberté a permis une première amélioration significative de l'affinité simulée, confirmée par RMN. L'étude de l'*iso*-propylcatéchol nous a aussi donné des résultats très prometteurs. Nous avons déterminé une affinité pour hPRDX5 de $\Delta G^{\circ}_b = -6.6 \pm 0.2 \text{ kcal.mol}^{-1}$ similaire à celle du 4NPC, et le ligand étant plus petit (11 atomes lourds contre 13 pour le 4NPC), cela donne lieu à une LE améliorée en comparaison du 4NPC, et significativement meilleure que celle du catéchol (voir chapitre I, p.42):

$$LE_{\text{CAT}} = 3.0/8 = 0.38 < LE_{\text{4NPC}} = 6.1/13 = 0.47 < LE_{\text{IPC}} = 6.6/11 = 0.60$$

Néanmoins vu la difficulté de synthèse de ce fragment nous avons préféré nous baser sur la structure du 4NPC pour la suite de la conception.

Tableau 10. Bilan de toutes les molécules testées par FM dont l'affinité pour hPRDX5 à été calculée

Ligand	Temps de simulation (ns)	Dernières X ns pour calcul affinité	ΔG°_b (kcal/mol)	K_D (mM)
catéchol	500	100	- 3.0 ± 0.2	6.9 ± 2.1
4-méthylcatéchol	500	100	- 4.2 ± 0.3	0.9 ± 0.4
2-phénylpropan-2-ol	500	100	- 1.9 ± 0.3	36 ± 9
benzène-1,2-diméthanol	300	50	- 2.3 ± 0.1	22 ± 6
gallate d'éthyle	500	100	- 5.6 ± 1.1	30 ± 50
cyclohexanediol-cis	300	150	- 2.8 ± 0.5	12 ± 9
cyclohexanediol-trans	500	100	- 2.9 ± 0.5	12 ± 11
acide (3,4-dihydroxyphényl) acétique	500	150	- 5.8 ± 1.1	0.3 ± 0.9
4-aminométhylcatéchol	500	150	- 2.1 ± 0.6	40 ± 44
4-éthylcatéchol	500	150	- 3.6 ± 0.7	3 ± 2
4- <i>iso</i> -propylcatéchol	600	120	- 6.6 ± 1.0	0.02 ± 0.02
4- <i>iso</i> -butylcatéchol	500	150	- 4.3 ± 0.7	0.8 ± 0.6
4- <i>tert</i> -butylcatéchol	500	100	- 5.5 ± 0.9	0.11 ± 0.09
4- <i>néo</i> -pentylcatéchol	1000	400	- 6.1 ± 0.8	0.06 ± 0.06
6-(<i>néo</i> -pentyl)biphényl-3,4-diol	600	200	- 6.9 ± 0.4	0.01 ± 0.01
4-benzyl-5-(<i>néo</i> -pentyl)catéchol	600	200	- 6.3 ± 0.8	0.06 ± 0.06
4-(<i>néo</i> -pentyl)-5-(phénylamino)catéchol	1400	500	- 8.2 ± 0.8	0.002 ± 0.001
4-[(4-aminophényl)amino]-5-(<i>néo</i> -pentyl)catéchol	500	100	- 5.7 ± 0.5	0.11 ± 0.14
4-[[2-(<i>néo</i> -pentyl)-4,5-dihydroxyphényl]amino]benzoate	500	100	- 4.8 ± 0.2	0.33 ± 0.15
(4-[[2-(<i>néo</i> -pentyl)-4,5-dihydroxyphényl]amino]phényl)méthanaminium	700	200	- 8.2 ± 1.0	0.003 ± 0.003

ΔG°_b = Valeur de l'énergie libre standard de liaison pondérée, prise sur les X dernières ns de simulation.

K_D = Valeur de la constante de dissociation du complexe PL pondérée au cours du temps, prise sur les X dernières ns

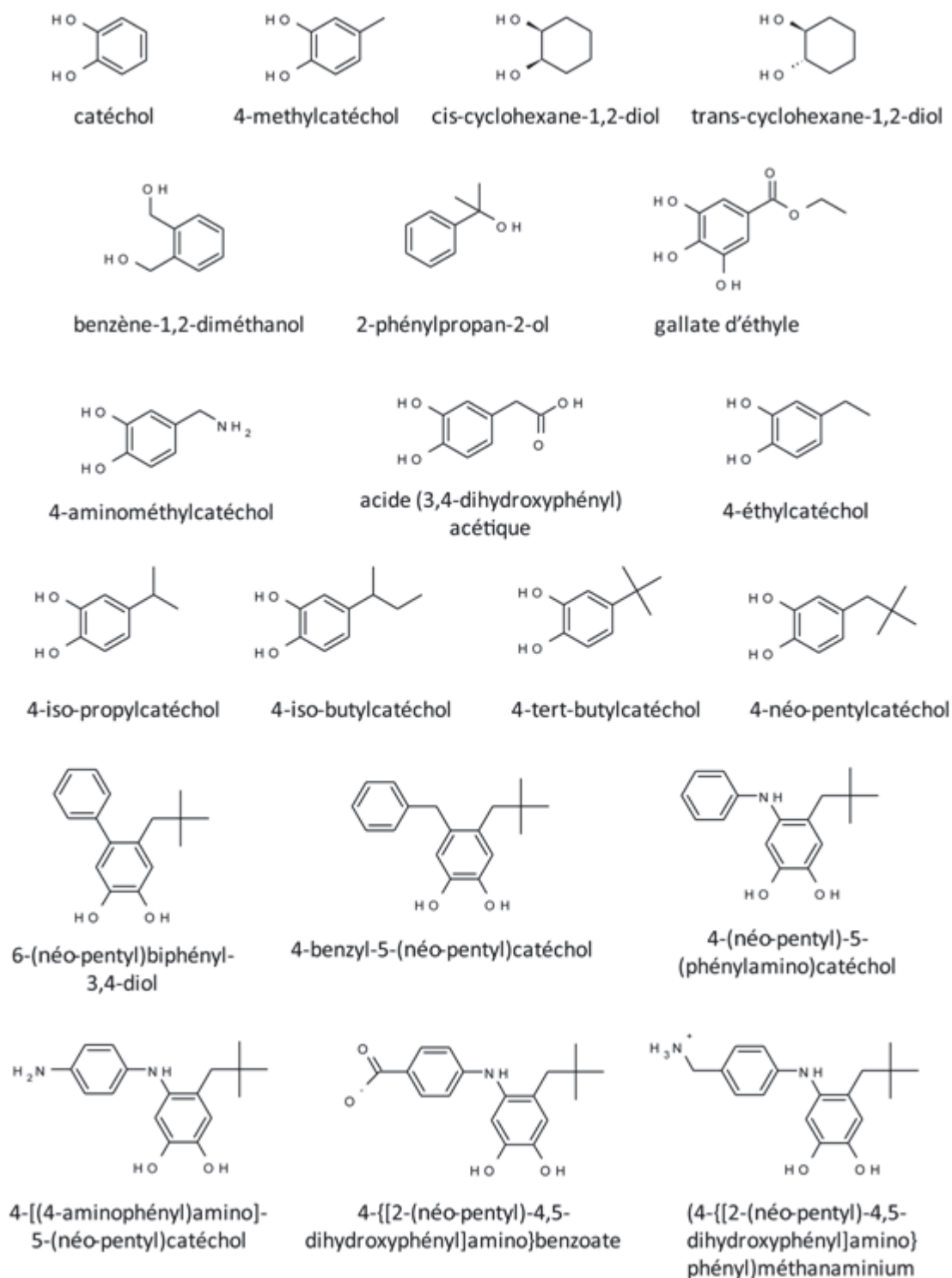


Figure 85. Structures des ligands dont l'affinité pour hPRDX5 a pu être estimée par FM lors de cette étude. Structures des fragments étudiés, et leur dénomination correspondante.

Ayant optimisé au mieux les interactions possibles pour cette molécule fragment, nous avons ensuite augmenté sa structure tridimensionnelle pour tenter de créer de nouvelles interactions avec hPRDX5. Ainsi, l'ajout d'un second substituant en position 5 du catéchol a été étudié dans le but d'augmenter la complexité du ligand tout en maintenant une affinité optimale avec hPRDX5. Le 4N5PC et son dérivé cationique Hit 3 ont finalement montré les meilleurs résultats d'affinité pour hPRDX5 par FM.

Grâce aux simulations de FM de ces différents composés avec hPRDX5, nous avons aussi pu mettre en évidence certaines caractéristiques de l'interaction. Nous avons notamment pu identifier l'échange conformationnel existant entre les complexes liés par une ou deux liaisons H avec la Cys47, et étudier la proportion de ces deux formes liées l'une par rapport à l'autre. L'interaction avec l'acide aminé Glu18 (chaîne B) qui dans certains cas tend à délocaliser l'interaction hors du site actif, pourrait potentiellement être une étape possible du chemin réactionnel. Elle a été décelée dans de nombreuses simulations et n'aurait pu être détectée uniquement par RMN. En effet, le passage par le complexe lié à Glu18 (chaîne B) dans le cas du ligand 4N5PC menait à sa liaison avec Cys47 *via* un changement conformationnel du ligand. Glu18 (chaîne B) est un acide aminé situé en surface, facilement accessible, et permettrait de positionner le ligand dans une conformation favorable à l'interaction avec Cys47 par deux liaisons H. Glu18 (chaîne B) n'a par contre pas été observé comme impliqué lors de la dissociation du complexe.

Enfin, la MD nous a permis d'observer en détail la dynamique du site actif de hPRDX5. Nous avons pu voir l'effet de l'interaction de différents ligands sur la compression ou l'élargissement de la cavité, et avons mis en évidence une possible corrélation entre ce paramètre et la stabilité du complexe sous sa forme liée. La FM nous a donc permis d'étudier à l'échelle microscopique l'affinité et l'interaction de nombreux ligands dérivés du catéchol avec hPRDX5 de manière quantitative, et nous a donné accès à de nombreux détails sur l'interaction PL, le ou les modes de liaison prédominants, et sur la dynamique interne du récepteur.

II. 4. 1 Comparaison avec le docking

Au début de ce manuscrit, j'ai mis en évidence quelques différences entre les simulations effectuées par MD et celles effectuées par *docking*. La plus importante d'entre elle est l'aspect dynamique global qui n'est pas présent dans les expériences de *docking* dans lesquelles la protéine est modélisée avec un squelette fixe et éventuellement les chaînes latérales flexibles.⁹⁶

Nous pouvons nous demander quel serait l'impact des différences entre le *docking* et la FM sur l'affinité estimée des ligands. Des études préliminaires par *docking* (AutoDock4.2.3) avaient déjà été effectuées au laboratoire sur les ligands CAT, 4MEC et 4TBC pour estimer leur mode de liaison dans

le cristal.^{91; 182}. Nous avons effectué l'analyse par *docking* moléculaire, de certains dérivés catéchols étudiés dans ce manuscrit pour hPRDX5. L'analyse par *docking* a été effectuée avec AutoDock Vina.⁹⁶ La protéine est modélisée de manière non solvatée à partir des données cristallographiques (3MNG.pdb), sans les atomes d'hydrogène, et les chaînes latérales des acides aminés sont modélisées avec leur flexibilité, excepté pour les acides aminés Pro, Ala, et Gly (qui ne possède qu'un atome d'hydrogène pour chaîne latérale). Les ligands sont aussi modélisés sans les atomes d'hydrogènes. Pour chacun des ligands, les 10 meilleures poses trouvées donnent lieu à une énergie empirique de liaison qui repose sur la somme des contributions des interactions inter et intramoléculaires formées par le ligand dans une fonction de scoring propre à AutoDock.⁹⁶ Les résultats obtenus sont visibles dans le Tableau 11 ci-dessous en prenant en compte l'affinité estimée de la meilleure des dix poses trouvées pour le *docking*.

Tableau 11. Résultats des énergies de liaison obtenues par *docking* pour les ligands CAT, 4MEC, 4TBC, 4NPC, Hit4 et Hit4N en complexe avec hPRDX5.

Comparaison avec les résultats obtenus par FM et RMN sur ces mêmes ligands

Ligand	ΔG_b° déterminée par RMN (kcal/mol)	ΔG_b° estimée par FM (kcal/mol)	ΔG_b estimée par docking (kcal/mol)
CAT	- 3.2 ± 0.1	- 3.0 ± 0.2	- 4.3
4MEC	- 4.1 ± 0.1	- 4.2 ± 0.3	- 4.9
4TBC	- 5.1 ± 0.1	- 5.5 ± 0.9	- 5.1
4NPC	- 6.6 ± 0.2	- 6.1 ± 0.8	- 5.3
4-benzyl-5-(néo-pentyl) catéchol	-	- 6.3 ± 0.8	- 5.9
4-(néo-pentyl)-5- (phenylamino)catéchol	-	- 8.2 ± 1.0	- 5.5

A première vue, on retrouve un classement similaire à celui obtenu par FM ou RMN au niveau de l'affinité des ligands étudiés pour hPRDX5. Cependant, les valeurs d'énergie obtenues sont différentes de celles estimées par FM. La différence d'énergie entre le catéchol et le 4N5PC mesurée par FM est de 5.2 kcal.mol⁻¹ tandis qu'elle n'est que de 1.2 kcal.mol⁻¹ par *docking*. Au vu des résultats obtenus, l'énergie de liaison semble d'autant plus sous estimée par *docking*, que la structure tridimensionnelle du ligand se rapproche du type « drug-like ». Cela peut être dû au fait que lors du *docking* la protéine est modélisée en gardant fixe le squelette carboné, contrairement à la FM où la

protéine est laissée libre, en ne contraignant que sa diffusion au sein de la boîte d'eau. La fixation du squelette influe sur la dynamique intrinsèque de la protéine, et il est possible que certaines interactions soient minimisées puisque la distance entre les atomes est prise en compte dans le calcul de la fonction de *scoring* et donc de l'énergie libre de liaison calculée. Par exemple, la contribution des interactions hydrophobes formées avec l'hélice Leu116 - Phe120 bien que prises en compte en *docking*, peuvent être sous estimées si l'hélice est positionnée plus loin du site actif dans les poses trouvées en *docking*. En FM, l'énergie étant calculée à chaque étape de simulation (4 fs) en prenant en compte la position de tous les atomes et en laissant libre le squelette carboné et les chaînes latérales de la protéine, la contribution de chaque interaction reflétera bien mieux la réalité. Notons que ce reflet de la « réalité » est aussi principalement dû au champ de force empirique utilisé (AMBER99SBildn ici).

II. 4. 2 Spécificité conformationnelle des différents ligands

Nous avons mis en évidence par FM des ligands d'affinité variable pour hPRDX5. Que ce soit en RMN ou en FM, l'énergie libre de liaison mesurée qui rend compte de l'affinité du ligand pour la protéine d'intérêt, est une moyenne des énergies des différentes conformations de liaison. Si la FM permet d'avoir une statistique de l'interaction par le biais des surfaces d'énergie libre et de la visualisation des clusters, la RMN elle, permet de localiser l'interaction par le biais des CSPs ou de « visualiser » des modifications conformationnelles.

En combinant ces observations, il est donc possible de mettre en avant la spécificité de certains ligands par rapport à d'autres. En RMN, si on analyse les CSPs induites par différents ligands, il est possible de comparer l'effet de l'interaction sur les acides aminés de la protéine d'intérêt. Sur l'histogramme donné en Figure 86, on peut comparer les CSPs induites par les ligands CAT, 4MEC, 4TBC et 4NPC sur hPRDX5, sur tous les acides aminés de la protéine. Ayant vérifié pour chaque expérience que hPRDX5 n'a pas subi de modification conformationnelle, cet histogramme rend compte de la quantité de CSP induite par ces 4 ligands sur chaque acide aminé ce qui permet de localiser l'interaction, et d'estimer leur spécificité.

Par exemple, on voit que le ligand CAT induit un maximum de perturbation de 0.1 ppm au niveau du site actif, mais perturbe aussi faiblement tous les acides aminés de la protéine. Si l'on regarde maintenant le ligand 4MEC, on voit que son maximum de perturbation ne dépasse pas 0.05 ppm, et que son interaction sur les autres acides aminés de hPRDX5 est faible mais non nulle, comme dans le cas du CAT. L'interaction avec les ligands 4TBC et 4NPC perturbent les acides aminés du site actif à hauteur de 0.15 ppm et 0.25 ppm respectivement, ainsi que les acides aminés de l'hélice Leu116 – Phe120 et les acides aminés 145 à 148 d'environ 0.1 ppm. Mais au contraire des ligands CAT et 4MEC, proportionnellement, ils n'auront qu'une influence minime sur les acides aminés non

impliqués dans l'interaction. Au final, les CSPs induites par ces 4 ligands sur des résidus non ou peu affectés par l'interaction sont équivalentes, mais diffèrent plus significativement pour les acides aminés impliqués dans l'interaction. Ainsi, on voit que le 4NPC possède une meilleure « spécificité conformationnelle » (un conformère dominant) que les autres dérivés car la différence des CSPs induites par son interaction avec hPRDX5 entre les acides aminés participant ou non à l'interaction est plus importante.

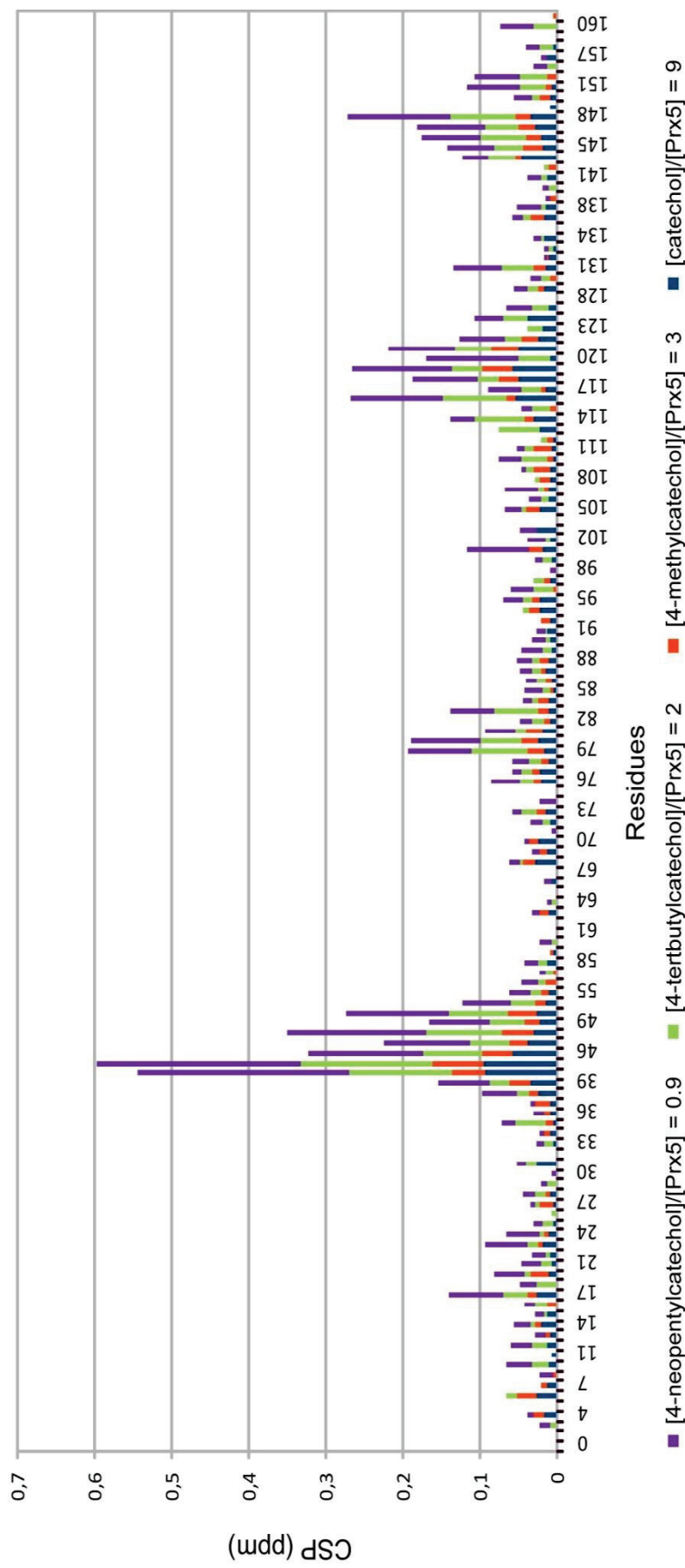


Figure 86. Histogramme empilé représentant la CSP totale subie par hPRDX5 sous l'influence des ligands catéchols CAT, 4MEC, 4TBC et 4NPC pour des concentrations en ligand et protéine équivalentes à 67% de la saturation totale.

Valeur de CSP correspondant à la somme des CSPs induites par CAT, 4MEC, 4TBC et 4NPC sur chaque acide aminé de la protéine, et visualisation de la contribution de chacun des ligands sur la perturbation totale. Résidus signalés par leur place dans la séquence protéique. Perturbation déterminée pour une concentration en ligand relative obtenue pour une valeur de $2K_D$, soit un rapport de concentration ligand sur protéine équivalent à 67% de la saturation.

III. ETUDE DE L'INHIBITION DE HPRDX5 PAR CINETIQUE ENZYMATIQUE

Jusqu'à présent, nous nous étions concentrés sur la capacité à reproduire par le calcul les mesures expérimentales des constantes d'équilibre d'affinité entre hPRDX5 et des ligands se liant dans le site actif de la forme réduite de l'enzyme. Logiquement, cette affinité devrait se traduire par une inhibition réversible et dépendante de la concentration du ligand, éventuellement inhibiteur (I). En inhibition enzymatique l'affinité est intégrée dans un système plus complexe que les mesures effectuées précédemment à l'équilibre thermodynamique. En effet, dans un modèle 1:1, nous avons l'équilibre **41** suivant :



En mesure de cinétique enzymatique, en considérant les inhibitions réversibles, on retrouve l'affinité dans au moins deux des états de l'enzyme : la forme libre de l'enzyme (E) et la forme complexée au substrat ou complexe de Michaelis (ES) (Figure 87).

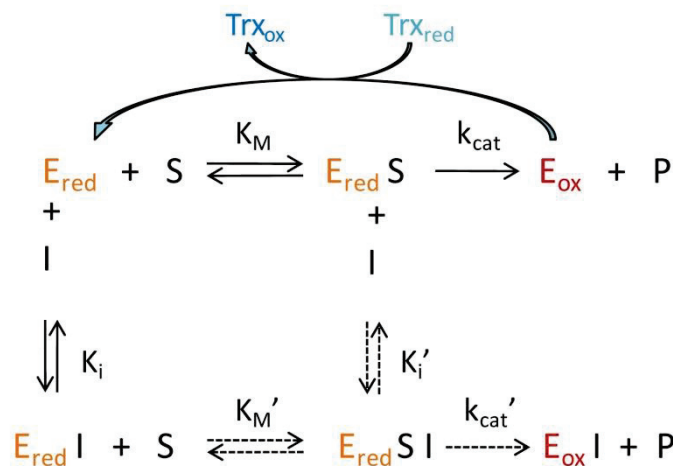


Figure 87. Mécanisme général de l'inhibition réversible des hPRDX.

La protéine est initialement sous forme réduite E_{red} et, par réaction avec le substrat S , va se retrouver sous forme oxydée E_{ox} lors de la formation du produit P . Cette forme oxydée de type acide sulfénique $-SOH$ est d'abord spontanément convertie en forme oxydée de type disulfure $-S-S-$ comme décrit en Figure 28 p.57, et est ensuite réduite de nouveau par une seconde enzyme la thioredoxine (Trx), selon le schéma visible Figure 27 p.56. Mécanisme pouvant être compétitif, incompétitif ou non-compétitif selon l'affinité de l'inhibiteur pour l'enzyme libre par rapport au complexe enzyme-substrat comme décrit en Figure 34 p.71.

Comme énoncé dans les buts de notre étude (p.1 et 2), l'objectif final était de montrer que nous avons un effet significatif sur l'activité de la protéine d'intérêt : hPRDX5. Ainsi les dérivés catéchols, dont les affinités pour hPRDX5 ont pu être estimées par FM précédemment dans ce manuscrit, devraient se comporter comme des inhibiteurs réversibles de hPRDX5. L'isoforme 5 des hPRDX possède un motif structural absent des autres isoformes de hPRDX : l'hélice Leu116 – Phe120. L'inhibition de ces dérivés catéchols pourrait alors avoir une certaine sélectivité d'action par rapport aux autres isoformes hPRDX puisque nous avons montré puis optimisé l'implication de cette hélice Leu116 - Phe120 avec certains ligands de hPRDX5. De plus, l'aspect prédictif offert par la FM trouve ici une application dans la détermination du mécanisme d'action des catéchols sur hPRDX5 si nous pouvions calculer l'affinité des ligands pour le complexe de Michaelis *ES*. La prédiction par FM de l'affinité déterminée à l'équilibre peut être mise en rapport indirect avec l'affinité impliquée dans la constante d'inhibition K_i (Figure 87). L'équilibre d'affinité sera déplacé par la conversion enzymatique (k_{cat}) et les autres équilibres avec le substrat ¹²⁴ et le complexe de Michaelis *ES* (K_i'). Si nous n'intégrons pas ici les phases réactives de formation de liaison chimique abordées dans les approches de type mécanique quantique (QM), nous pourrions cependant intervenir sur la prédiction des équilibres définis par les constantes K_i , K_M , K_i' .

III. 1. 1 Etude de l'interaction de hPRDX5 avec son substrat

Avant d'étudier l'inhibition des différents dérivés catéchols, nous avons cherché à déterminer l'affinité moyenne pour hPRDX5 de H₂O₂, son substrat naturel.

La simulation de FM pour le système comprenant H₂O₂ et hPRDX5 a été lancée sur 200 ns et a donné lieu à 11 évènements de très courte durée (< 1 ns). A la fin de la simulation, la différence d'énergie libre de liaison du complexe converge vers une valeur proche de 0 indiquant un processus ni favorable ($\Delta G_b^\circ < 0$) ni défavorable ($\Delta G_b^\circ > 0$). Du fait de sa très petite taille, H₂O₂ diffuse très rapidement et librement comme le solvant (H₂O) dans le site actif. En cartographiant l'énergie on peut cependant observer deux bassins énergétiques correspondant à l'interaction dans le site actif avec Cys47, correspondant dans la simulation à un complexe *ES*. La faible différence entre l'état libre et lié est en rapport avec l'absence d'affinité apparente de H₂O₂ pour hPRDX5. Durant cette simulation, n'ont été observés que des évènements de liaison mettant en jeu une seule liaison H entre H₂O₂ et hPRDX5, avec un angle SG-O_A-O_B de 50° (ou de 100° du fait de la symétrie de la molécule, Figure 88).

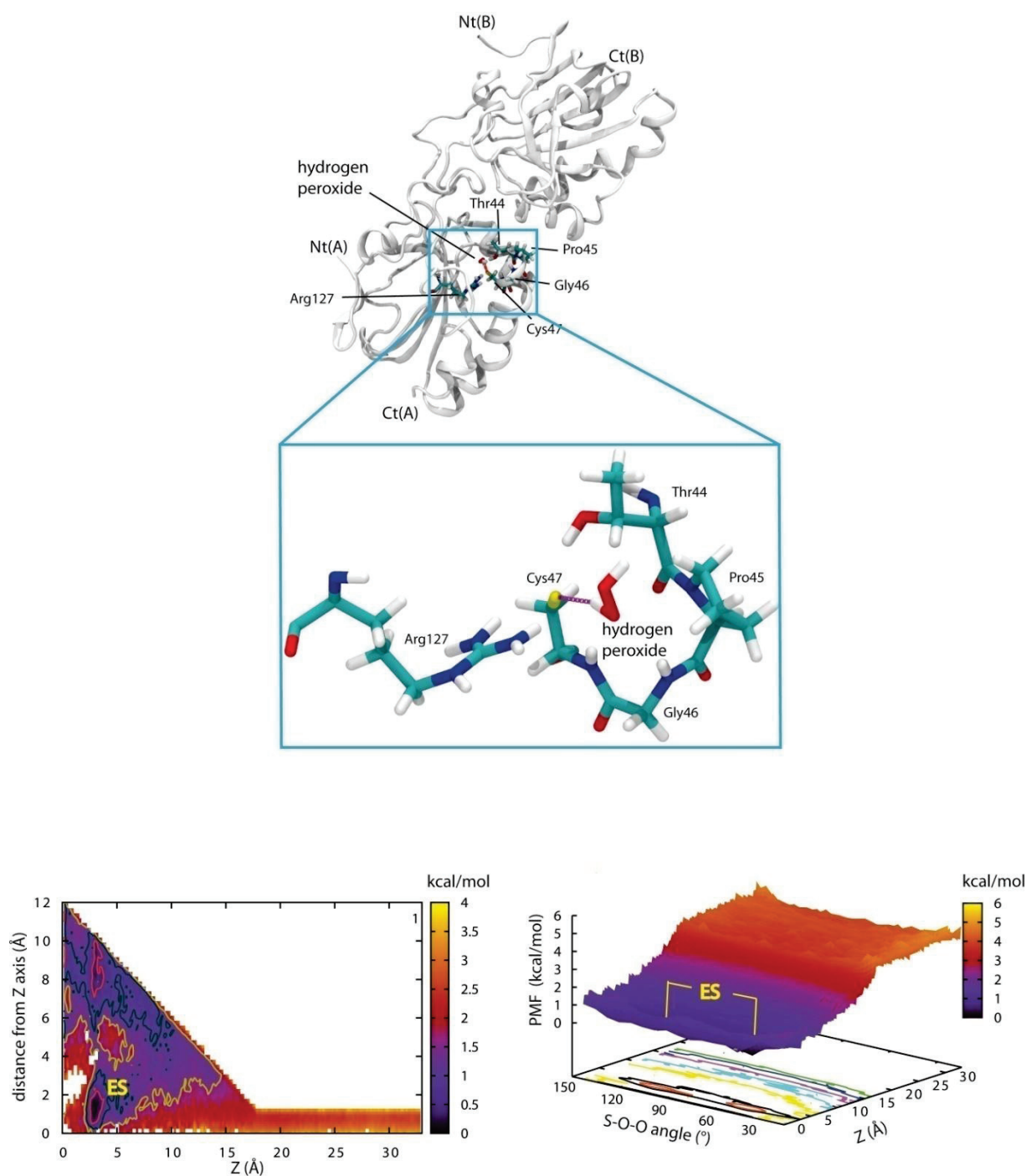


Figure 88. Simulation du peroxyde d'hydrogène avec hPRDX5 : Energie du complexe apparent ES.

A) Complexe de Michaelis apparent (*ES*) formé par hPRDX5 et H_2O_2 , durant la simulation de FM avec les principaux acides aminés du site actif représentés en bâtons (Thr44, Pro45, Gly46, Arg127 et Cys47), et les *N-term* et *C-term* de chaque chaîne (A et B) indiqués Nt et Ct. La liaison hydrogène entre le thiolate de Cys47 et H_2O_2 est visible en ligne pointillée violette sur l'agrandissement. B) FES 2D (CV1-CV2) et 3D (CV1-CV3). Cartographie de l'énergie ($kcal \cdot mol^{-1}$) utilisant des contours iso-énergétiques avec un intervalle de $0.5 kcal \cdot mol^{-1}$.

III. 1. 2 Etude de l'inhibition de hPRDX5 par les dérivés catéchols

Les dérivés catéchols ayant une affinité pour le site actif de hPRDX5, devraient se comporter comme des inhibiteurs réversibles de hPRDX5 en l'absence de réaction entre catéchols et hPRDX5. L'amélioration de l'affinité qui a pu être mise en évidence sur les dérivés catéchols peut-elle être corrélée à une amélioration de l'inhibition de hPRDX5 ? Et ces dérivés inhibent-ils hPRDX5 d'après le même mécanisme ? Dans le but de vérifier l'inhibition réversible de hPRDX5 par les dérivés catéchols, nous avons donc effectué des expériences d'inhibition et de mesure d' IC_{50} , par cinétique enzymatique.

- Inhibition de hPRDX5 par le dérivé 4NPC, mesure d' IC_{50}

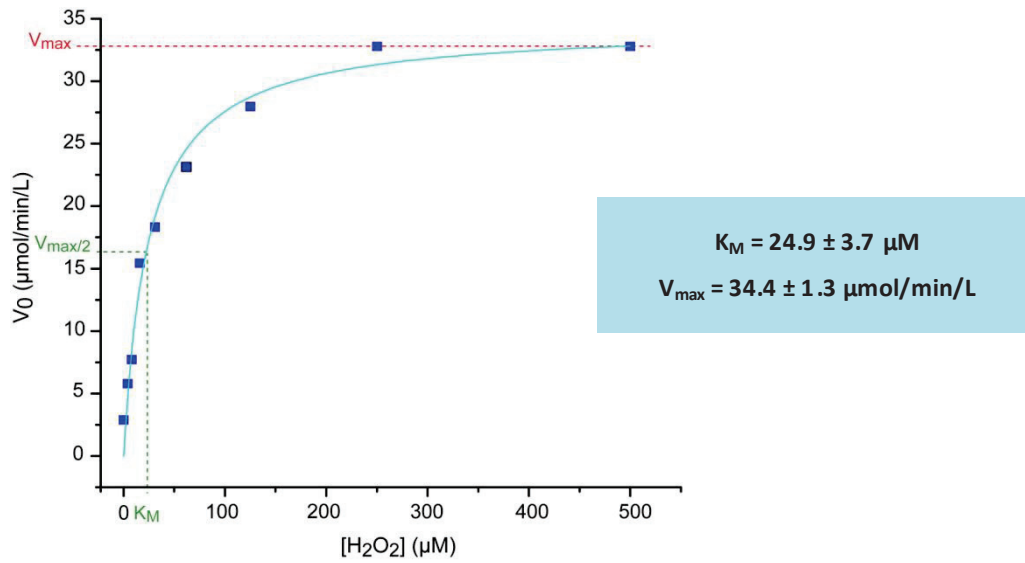
Une composition optimale du système protéique avait été proposée par Melissa Chow dans son travail de thèse au sein du laboratoire, décrit en chapitre II (p.96).¹⁸⁸. Afin de déterminer l' IC_{50} du 4NPC, il a tout d'abord fallu vérifier la valeur du K_M relatif à la réaction de hPRDX5 avec H_2O_2 pour ce système (déterminé à $24.2 \pm 3.2 \mu M$). La vitesse initiale de réaction est analysée pour différentes concentrations en H_2O_2 allant de 0 à 500 μM , et à concentration constante en hPRDX5. La mesure de cette constante de Michaelis K_M , associée à une valeur de vitesse maximum de réaction V_{max} , permet de contrôler l'activité enzymatique du système, et l'activité spécifique de hPRDX5.

Dans le cas des lots de protéines hPRDX5, γTrx et $\gamma TrxR$ que j'ai produites, on obtient des valeurs de $K_M = 24.9 \pm 3.7 \mu M$ et $V_{max} = 34.4 \pm 1.3 \mu mol \cdot min^{-1} \cdot L^{-1}$ (Figure 89) en accord avec les résultats précédemment obtenus. Pour déterminer la valeur de l' IC_{50} , la vitesse initiale de réaction V_0 est mesurée indirectement par la consommation du NADPH suivie à 340 nm en utilisant une concentration en H_2O_2 égale au K_M et une concentration variable en ligand comme décrit en chapitre II (p.96). En normalisant les vitesses V_0 ($\mu mol \cdot min^{-1} \cdot L^{-1}$) par rapport à la plus haute valeur trouvée, on exprime le pourcentage d'activité résiduelle en fonction de la concentration en 4NPC, et on obtient un IC_{50} relatif de $40 \pm 5 \mu M$. On peut noter également que l'activité de hPRDX5 n'est pas complètement inhibée au vu des 27% d'activité résiduelle observés (Figure 89.B).

- Mécanisme de l'inhibition de hPRDX5 par le dérivé 4NPC

Afin de mieux comprendre l'inhibition de la protéine par le 4NPC, nous avons alors voulu déterminer le mécanisme d'inhibition mis en jeu. Pour étudier ce mécanisme, V_0 a été mesurée en fonction de la concentration en substrat pour 5 concentrations en inhibiteur. Les courbes obtenues ont été analysées via le logiciel DynaFit afin de déterminer le mécanisme d'inhibition (voir annexe 6).

A)



B)

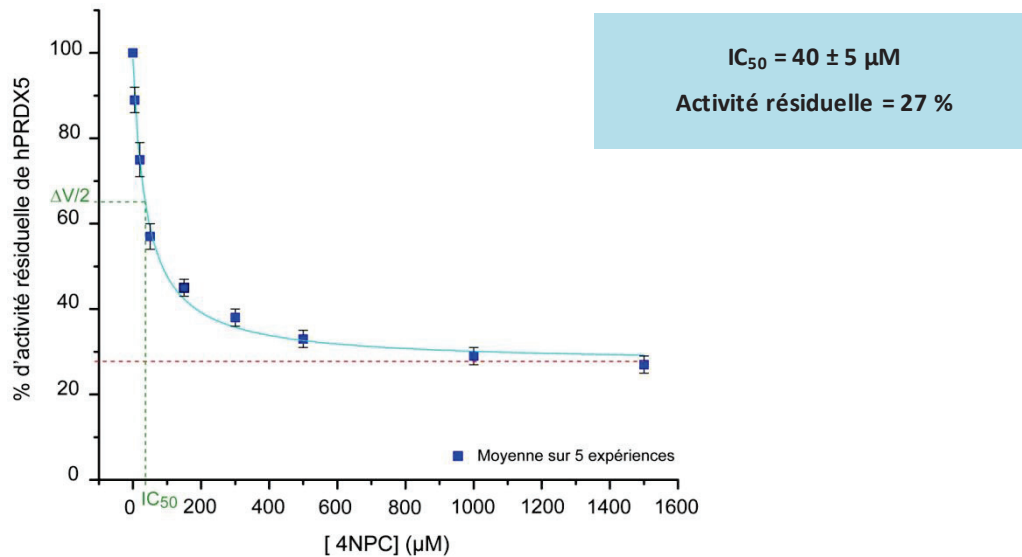
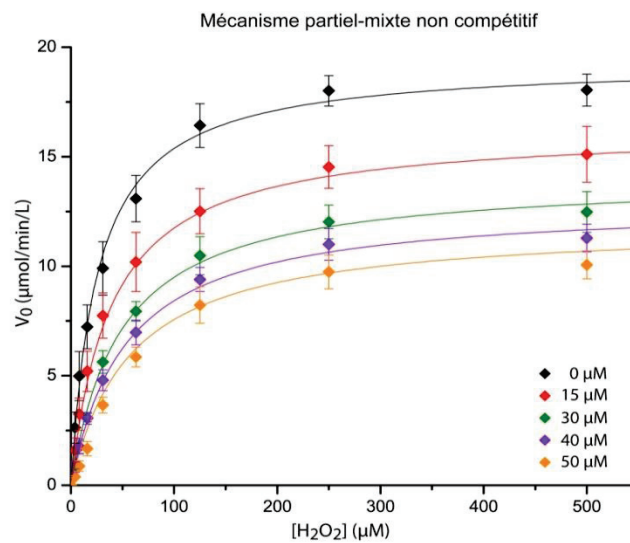


Figure 89. Etude de l'inhibition du 4NPC pour hPRDX5.

A) $V_0 = f([\text{H}_2\text{O}_2])$. Vitesse initiale de la réaction en $\mu\text{mol} \cdot \text{min}^{-1} \cdot \text{L}^{-1}$. Correspondance avec une courbe théorique de type : $y = A \times x / B + x$ pour la détermination de K_M et V_{max} . B) Courbe de détermination de l' IC_{50} du 4NPC : % d'activité résiduelle de hPRDX5 en fonction de la concentration en 4NPC introduite. Points expérimentaux représentés comme moyenne de 5 expériences, erreur sur ces points calculée comme l'erreur standard sur les 5 mesures indépendantes. Correspondance avec la courbe théorique de type $y = C + (B \cdot A/x + A)$ pour le calcul de l' IC_{50} (constante A) et de l'activité résiduelle (constante C).

Les courbes expérimentales sont corrélées avec des courbes théoriques de type hyperbole, et par analyse de la variation des valeurs de V_{max} et de K_M en fonction des différentes concentrations en inhibiteur, il est possible de déterminer le mécanisme d'inhibition (Figure 90). De plus, le logiciel permet l'obtention des valeurs des constantes cinétiques K_i , K_i' , k_{cat} et k_{cat}' calculées selon le type de mécanisme. Dans le cas du 4NPC, l'analyse des résultats obtenus indique un mécanisme d'inhibition partiel-mixte non compétitif (Figure 34 p.71). Dans ce modèle d'inhibition, le 4NPC peut se lier à l'enzyme libre E ou au complexe enzyme substrat ES avec deux affinités différentes respectivement K_i et K_i' .

A)



B)

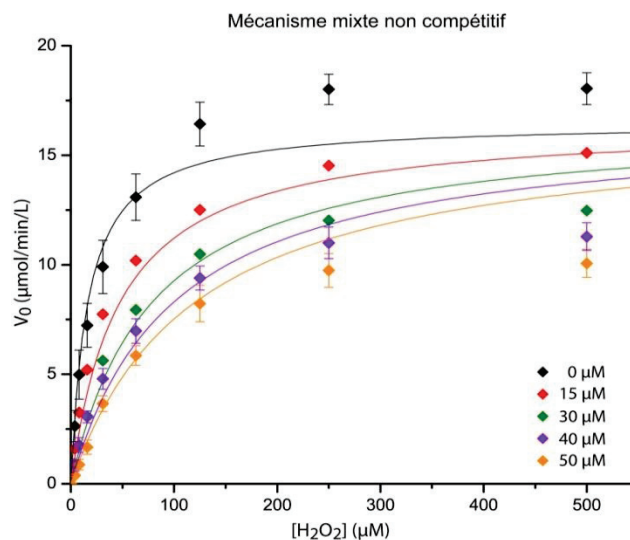


Figure 90. Analyse du mécanisme d'inhibition de hPRDX5 par le 4NPC.

Détermination du modèle d'inhibition enzymatique. Vitesse de réaction initiale en $\mu\text{mol}\cdot\text{min}^{-1}\cdot\text{L}^{-1}$ en fonction de la concentration en substrat H_2O_2 . Points expérimentaux représentés comme moyenne de 3 expériences indépendantes. L'erreur sur ces points est calculée comme l'erreur standard sur les 3 mesures. La correspondance avec les modèles théoriques calculés par DynaFit est représentée en traits pleins. A) Mécanisme partiel mixte non compétitif. B) Mécanisme mixte non compétitif.

Les valeurs des différentes constantes relatives à ce modèle d'inhibition, déterminées par DynaFit, pour ces expériences avec 4NPC sont les suivantes : $K_s = 26 \pm 2 \mu\text{M}$ ($= K_M$ voir p.67) ; $K_i = 19 \pm 5 \mu\text{M}$; $K_i' = 80 \pm 40 \mu\text{M}$; $k_{cat} = 2.15 \pm 0.05 \text{ s}^{-1}$ et $k_{cat}' = 0 \sim 0.67 \text{ s}^{-1}$. On remarque tout d'abord que pour $0 \mu\text{M}$ de 4NPC, le V_{max} observé qui devrait correspondre à celui trouvé lors de la première expérience de cinétique enzymatique effectuée, est différent. En effet, ici $V_{max} \approx 18 \mu\text{mol}\cdot\text{min}^{-1}\cdot\text{L}^{-1}$ à $0 \mu\text{M}$ de 4NPC, tandis que $V_{max} \approx 35 \mu\text{mol}\cdot\text{min}^{-1}\cdot\text{L}^{-1}$ dans l'expérience initiale (cinétique de type Michaelis-Menten). Cela s'explique possiblement par l'utilisation de lots différents de protéines γTrx , γTrxR et hPRDX5. On note ensuite que la valeur de K_i' est supérieure à celle de K_i ce qui indique que 4NPC se lierait préférentiellement à l'enzyme libre (E_{red}) plutôt qu'au complexe enzyme-substrat ES . De plus, on peut voir que la valeur de k_{cat}' est calculée à 0, avec une grande erreur sur la mesure. Si k_{cat}' était nulle, le mécanisme d'inhibition serait alors mixte non compétitif, et non partiel. Si l'on vérifie maintenant la correspondance des courbes théoriques pour un modèle d'inhibition mixte non compétitif avec les résultats expérimentaux, on voit que ce modèle ne correspond pas à notre mécanisme d'inhibition (Figure 90). k_{cat}' est donc bien différent de 0, avec une limite haute de 0.67 s^{-1} , ce qui implique que même si elle est faible, il y a bien conversion de l'enzyme à partir du complexe ternaire ESI , de même qu'à partir du complexe ES puisque k_{cat} n'est pas nulle. Cela tend à expliquer l'activité résiduelle observée lors des mesures d' IC_{50} .

- Comparaison de l'inhibition des autres dérivés catéchols

Melissa Chow a étudié en parallèle de ce travail l'inhibition des hPRDXs et notamment l'inhibition de hPRDX5 par les dérivés CAT, 4MEC et 4TBC.¹⁸⁸ Les IC_{50} présentés en Figure 91 et regroupés dans le Tableau 12 montrent que le ligand qui inhibe le mieux hPRDX5 est le 4NPC.

Non seulement la valeur de l' IC_{50} est plus faible, mais la valeur du pourcentage d'activité résiduelle observé est aussi plus faible. Ainsi dans le cas du CAT, l'activité de la protéine ne peut être inhibée que de 60 % au maximum (40 % d'activité résiduelle), contre 75 % dans le cas du 4NPC (25 % d'activité résiduelle). Les résultats d'inhibition sont en bonne corrélation avec ceux obtenus dans le cas de l'affinité puisqu'on retrouve le même classement au niveau de ces ligands, mais aussi un ratio similaire pour le 4NPC par exemple dont l'affinité est 100 fois supérieure à celle du CAT, tout comme l'est le rapport de leurs IC_{50} .

De plus, les expériences ont montré que tous ces dérivés catéchols présentaient le même mécanisme d'inhibition pour hPRDX5: le mécanisme partiel-mixte non compétitif. Si l'on regarde de manière plus détaillée la valeur des différentes constantes d'inhibition associées pour l'inhibition de hPRDX5 par ces différents dérivés catéchols, on observe que la valeur des K_i' est toujours supérieure à la valeur des K_i , indiquant une liaison préférentielle de ces inhibiteurs avec l'enzyme libre plutôt qu'avec le complexe ES .

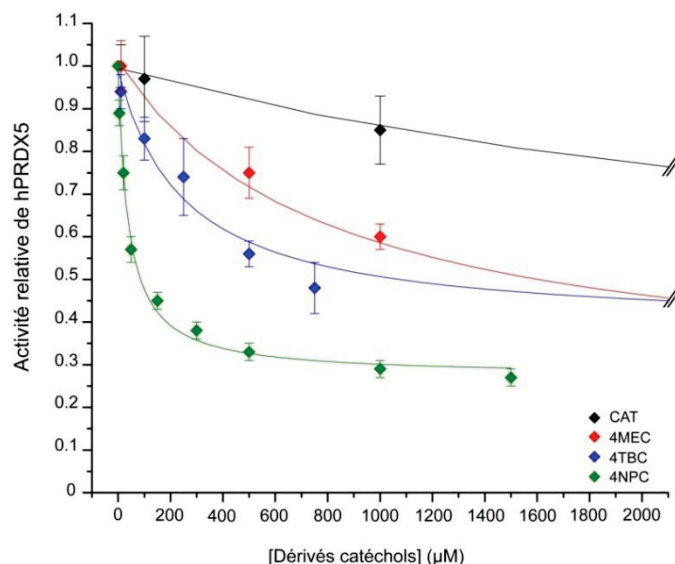


Figure 91. Comparaison des IC_{50} obtenus pour les différents dérivés catéchols étudiés avec hPRDX5.

Courbes de détermination des IC_{50} des différents ligands étudiés avec hPRDX5. Vitesse initiale de réaction normalisée par rapport à la plus haute valeur donnant l'activité relative de hPRDX5 exprimée en fonction de la concentration en ligand ajoutée. Le graphique est agrandi sur la région [0-2000 μM] permettant la comparaison avec 4NPC et ne montre que les premiers points pour CAT, 4MEC et 4TBC. Les gammes de concentrations en ligand sont adaptées en fonction de l'affinité : [CAT] = 0-75 mM ; [4MEC] = 0-15 mM ; [4TBC] = 0-5 mM et [4NPC] = 0-1.5 mM.

Tableau 12. Comparaison de l'inhibition de hPRDX5 par les ligands catéchols étudiés

Ligand	CAT [§]	4MEC [§]	4TBC [§]	4NPC
IC_{50} (mM)	3.73 ± 0.85	0.82 ± 0.14	0.25 ± 0.06	0.04 ± 0.01
Activité résiduelle (%)	40	35	35	25
Modèle	Partiel-Mixte Non compétitif			
K_s (μM)	40.7 ± 4.0	33.7 ± 4.6	36.9 ± 4.4	26 ± 2
K_i (mM)	1.81 ± 0.46	0.33 ± 0.13	0.09 ± 0.04	0.019 ± 0.005
K_i' (mM)	4.04 ± 1.13	0.36 ± 0.12	0.18 ± 0.10	0.080 ± 0.04
k_{cat} (s^{-1})	3.08 ± 0.08	3.11 ± 0.01	3.05 ± 0.10	2.15 ± 0.05
k_{cat}' (s^{-1})	≈ 0 $(5.10^{-4} \pm 0.4)$	1.07 ± 0.25	2.35 ± 0.18	≈ 0 $(5.10^{-4} \pm 0.67)$

[§] M. Chow, L. Troussicot et al. *Biochemistry*. 2016.¹⁸⁸

% d'activité résiduelle estimé graphiquement sur la moyenne de plusieurs expériences.

Différence des valeurs de K_s calculées estimées due aux lots différents de protéines utilisées.

Si l'on s'intéresse cette fois aux constantes catalytiques k_{cat} et k_{cat}' représentatives des deux chemins menant à la conversion de H_2O_2 et à l'oxydation de l'enzyme, on remarque que dans le cas des dérivés 4MEC et 4TBC la valeur de k_{cat}' est similaire à la valeur de k_{cat} . Au contraire, les dérivés CAT et 4NPC ont une constante catalytique k_{cat}' faible ce qui pourrait indiquer que ces ligands se rapprochent d'une inhibition mixte non compétitive, sans l'atteindre. Néanmoins, les valeurs de k_{cat}' pour ces deux ligands présentent une erreur significative importante (coefficient de variation $\approx 10^6$) et les données expérimentales obtenues ne s'ajustent pas avec un modèle d'inhibition non compétitive mixte (induisant $k_{cat}' = 0$), comme expliqué précédemment dans le cas du 4NPC. Les 4 dérivés inhibent donc hPRDX5 par un même mécanisme : partiel-mixte non compétitif avec une nette amélioration de l'activité résiduelle pour le 4NPC, en comparaison du dérivé CAT initial.

III. 1. 3 Estimation du mécanisme d'inhibition par FM

Nous avons vu que la FM permettait la prédiction des affinités protéine-ligand (protéine-inhibiteur dans le cas présent) de manière réaliste et précise. Dans le cas d'un système enzymatique, les affinités de l'inhibiteur pour l'enzyme libre et le complexe enzyme-substrat *ES* vont permettre la détermination du mécanisme d'inhibition de ces enzymes. La FM pourrait donc, en ce sens, devenir prédictive des mécanismes d'inhibition (réversible) impliqués dans le mode d'action de ces enzymes par le biais de la prédiction des affinités relatives aux constantes d'inhibition K_i et K_i' . Afin de valider ou d'infirmer cette hypothèse, nous avons alors utilisé la FM pour modéliser l'interaction de certains dérivés catéchols étudiés précédemment, avec le complexe *ES* (hPRDX5- H_2O_2) lié. Ces simulations ont été effectuées sur deux dérivés simples ayant montré une bonne affinité pour l'enzyme libre : 4TBC et 4NPC. Pour modéliser le complexe lié *ES*, une contrainte de distance de 2Å a été appliquée entre le substrat H_2O_2 et l'atome SG de Cys47, afin de limiter sa diffusion tout en lui permettant de s'adapter continuellement à la géométrie du site actif (voir partie I.4 du chapitre II p.79). Des simulations FM de 500 ns ont été effectuées pour chacun des deux ligands, menant aux affinités calculées répertoriées dans le Tableau 13 ci-dessous.

La simulation du 4TBC avec le complexe *ES* hPRDX5- H_2O_2 a donné lieu à 8 évènements de liaison ce qui a permis un bon échantillonnage des formes libres et liées du complexe. L'affinité estimée du 4TBC pour le complexe *ES* est plus faible de 1 kcal.mol⁻¹ par rapport à l'affinité estimée pour l'enzyme libre. La cartographie de l'énergie en fonction des variables CV1 (distance sur l'axe Z) et CV2 (distance par rapport à l'axe Z) des systèmes modélisés en absence et en présence de substrat nous montre que l'interaction PL est déplacée vers l'extérieur de l'entonnoir (Figure 92). On visualise toujours l'interaction avec Cys47 à 5 Å sur Z, mais celle-ci est plus faible en énergie que l'interaction avec Glu18 (chaîne B) à 9 Å sur Z qui semble avoir été renforcée.

Tableau 13. Comparaison des résultats d'affinité estimés par FM pour l'enzyme libre hPRDX5 et pour le complexe ES hPRDX5-H₂O₂ : Cas des ligands 4TBC et 4NPC.

Ligand	Affinité estimée pour hPRDX5 ΔG°_b (kcal/mol)	Affinité estimée pour le complexe ES ΔG°_b (kcal/mol)
4TBC [§]	- 5.5 ± 0.9	- 4.4 ± 0.5
4NPC	- 6.1 ± 0.8	- 3.5 ± 0.6

[§] M. Chow, L. Troussicot et al. *Biochemistry*. 2016.¹⁸⁸

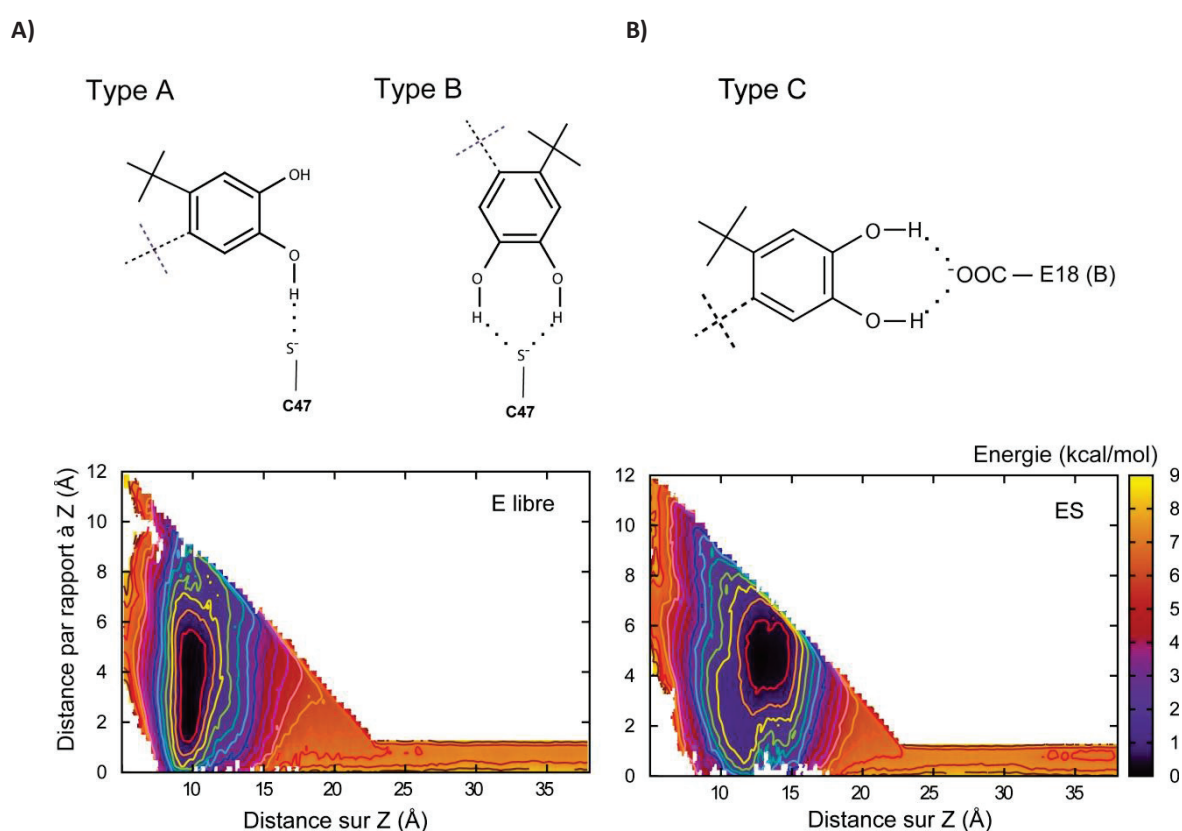


Figure 92. Représentation schématique des conformations de liaison (haut) et FES représentant l'énergie libre du complexe hPRDX5-4TBC en absence et en présence de H₂O₂ en fonction de CV1 et CV2 (bas).

FES en fonction de CV1 et CV2, et visualisation des conformations représentatives des bassins d'interaction observés dans le cas de l'interaction du 4TBC : A) en absence et B) en présence du substrat (H₂O₂) dans le site actif de hPRDX5. Les courbes de niveaux sont représentées avec un intervalle de 0.5 kcal.mol⁻¹.

Le déplacement de ce bassin dû à la présence du substrat nous indique que le ligand éprouve des difficultés à venir se lier à Cys47 en présence de H₂O₂, et qu'il aura alors plus de facilité à aller former des interactions de surface (comme observé avec Glu18 (chaîne B)). Lors de sa liaison avec Cys47, le

ligand tend à se lier par le biais d'un seul de ses groupes hydroxyles compte tenu de l'espace disponible (conformation A), mais la seconde conformation de liaison B, est tout de même observée. Lors de son interaction avec Glu18 (chaîne B) par contre, la seule conformation observée est celle mettant en jeu deux liaisons H (Figure 92.B).

Dans le cas du 4NPC, 4 évènements de liaison ont été observés en présence de H₂O₂, sensiblement plus courts en comparaison avec la simulation du 4NPC effectuée sur l'enzyme libre. De plus, une perte de plus de 2 kcal.mol⁻¹ est enregistrée pour l'affinité du 4NPC pour le complexe hPRDX5-H₂O₂ par rapport à celle calculée dans le cas de l'enzyme libre. Par visualisation de la trajectoire, on observe la coexistence des conformations de liaison à 1 et 2 liaisons H observées lors de l'interaction pour l'enzyme libre, avec dans le cas présent une prédominance de la conformation à 1 liaison H.

En cartographiant l'énergie libre du complexe en fonction des variables collectives CV1 et CV2, on distingue 2 bassins d'énergie correspondants respectivement à l'interaction avec Cys47, à 15 Å sur l'axe Z, et à l'interaction avec Glu18 (chaîne B), à 18 Å sur Z (Figure 93.A). Les courbes de niveau indiquent que ces deux bassins d'interaction sont de même énergie. La présence du substrat induit un déplacement de l'interaction de Cys47 vers Glu18 (chaîne B).

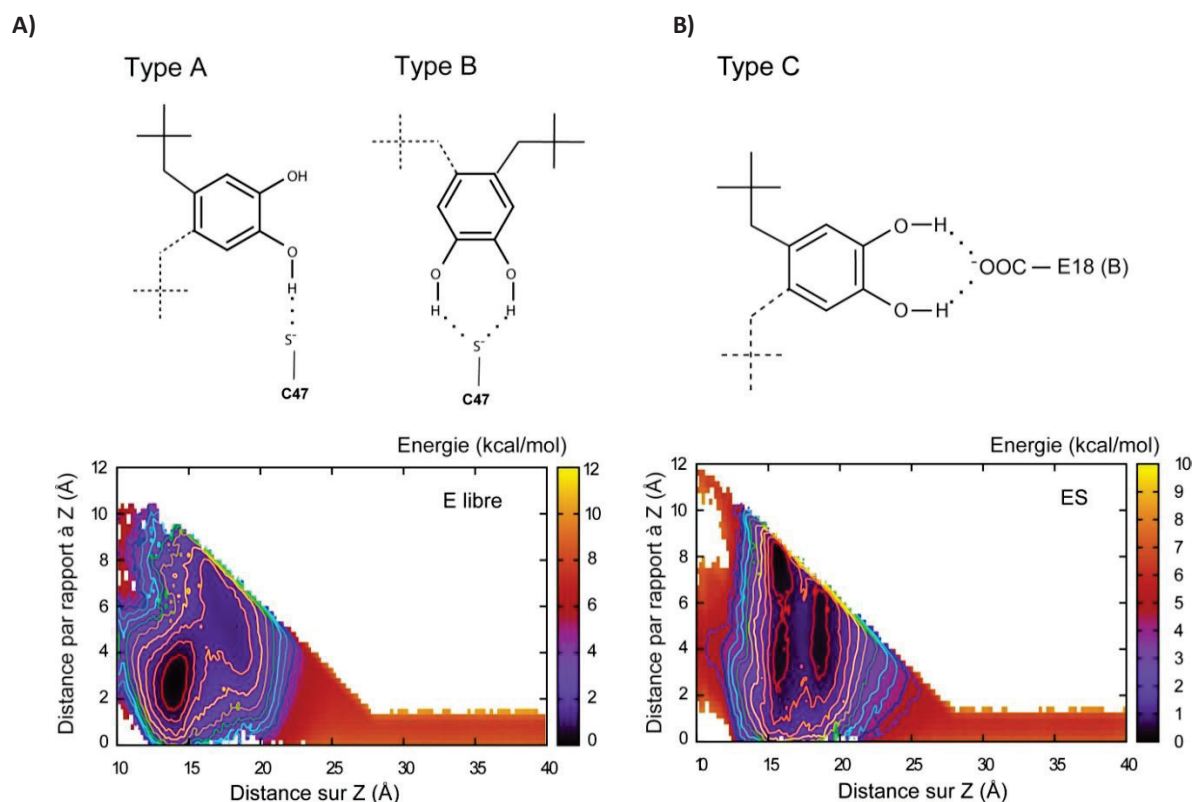


Figure 93. Représentation schématique des conformations de liaison (haut) et FES représentant l'énergie libre du complexe hPRDX5-4NPC en absence et en présence de H₂O₂ en fonction de CV1 et CV2 (bas).

FES en fonction de CV1 et CV2, et visualisation des conformations représentatives des bassins d'interaction observés dans le cas de l'interaction du 4NPC : A) en absence et B) en présence du substrat (H₂O₂) dans le site actif de hPRDX5. Les courbes de niveaux sont représentées avec un intervalle de 0.5 kcal.mol⁻¹.

Si l'on cartographie l'énergie du complexe hPRDX5-4NPC en fonction des variables CV1 et CV3 (angle SG-O_A-O_B), il n'y a aucune différenciation visible en termes d'angle, les bassins d'interaction sont définis entre 30° et 110° environ. Le complexe obtenu par l'interaction avec Glu18 (chaîne B) est par contre uniquement formé par le biais de 2 liaisons H comme visible sur la Figure 93.B. Par regroupement et analyse des différentes conformations de liaison observées lors de ces simulations, on voit que le ligand se lie à Cys47 en présence du substrat dans le site actif, par une ou deux liaisons H, avec prédominance de la conformation à une liaison H (Figures 92 et 93).

Le substrat étant déjà lié à Cys47, la place disponible pour la liaison du ligand est nettement plus faible. Malgré une perte d'affinité observée, la liaison de ces dérivés catéchols au complexe *ES* nous permet de pouvoir prédire un mécanisme d'inhibition non compétitif, que l'on peut ensuite corréler avec des expériences de cinétique enzymatique.

Nous avons donc obtenu par FM des affinités plus faibles pour le complexe *ES* que pour l'enzyme libre. Ces résultats sont en accord avec les valeurs des constantes d'inhibition K_i et K_i' mesurées par cinétique enzymatique et qui nous ont montré une liaison préférentielle de l'inhibiteur à l'enzyme libre. Plus précisément, dans le cas du dérivé 4TBC nous avons vu que la constante K_i' était similaire à K_i (rapport $K_i'/K_i \approx 2$). On peut relier ce rapport à la faible différence d'affinité observée pour hPRDX5 en présence et en absence de substrat (1 kcal.mol⁻¹). Au contraire, les constantes d'inhibition mesurées pour le 4NPC présentaient une différence deux fois plus importante ($K_i'/K_i \approx 4$) indiquant une plus nette préférence pour l'enzyme libre par rapport au complexe *ES*. On retrouve cette différence dans l'estimation des affinités pour hPRDX5 en absence et en présence de substrat avec une différence de 2.5 kcal.mol⁻¹.

III. 1. 4 Etude de la sélectivité des dérivés catéchols

hPRDX5 est l'une des 6 isoformes de hPRDX que l'on retrouve dans les cellules cérébrales.¹²⁰ Notre étude s'est tout d'abord portée sur hPRDX5 qui a été montré exister en homodimère symétrique stable, par rapport aux protéines hPRDX1 à 4 qui sont des oligomères de plus grande taille (tétramère à dodécamère) et à hPRDX6 qui ne possédait pas de structure cristallographique connue au début de ce projet.

Afin d'avoir une première idée de la sélectivité des dérivés catéchols, les construits nécessaires à l'expression chez *E. Coli* de hPRDX1 et hPRDX2 nous ont été confiés par B. Knoops (Université de Louvain, BEL). Ces protéines ont été produites au laboratoire par M. Chow lors de sa thèse, en suivant le même protocole de production que pour hPRDX5 (voir chapitre II.2). M. Chow a ensuite mesuré par chromatographie d'exclusion stérique les masses apparente de hPRDX1 et hPRDX2 qui sont de l'ordre de 300 kDa, correspondant à l'erreur près, à des oligomères décameriques ou

dodécamérique (monomères de 22kDa). Les mesures d'affinité des dérivés catéchols pour ces deux isoformes hPRDX1 et hPRDX2 seront effectuées par STD car leur haut poids moléculaire, dû à leur oligomérisation, rend difficile la détection par ^{15}N -HSQC et ne permet pas la mesure des CSPs.

- Etude de l'inhibition de hPRDX1 et hPRDX2 par le 4NPC

La valeur de $I'_{C_{50}}$ du 4NPC pour hPRDX1 et hPRDX2 a été déterminée comme précédemment. Pour cela, la valeur du K_M de ces deux protéines a été déterminée par M. Chow en utilisant une gamme en H_2O_2 allant de 0 à 25 μM . Ces expériences ont menées aux résultats suivants : K_M hPRDX1 = $2.1 \pm 0.4 \mu\text{M}$ et K_M hPRDX2 = $2.5 \pm 0.6 \mu\text{M}$. Les valeurs obtenues sont différentes d'un facteur 10 de celles déterminées dans le cas de hPRDX5. Les protéines hPRDX1 et hPRDX2 se comportent de manière similaire (K_M voisins), ce qui peut être expliqué par leur ressemblance structurale (voir III.1 du chapitre I). Le système utilisé pour les expériences de cinétique enzymatique dans le cas de hPRDX1 et hPRDX2 est similaire à celui utilisé dans le cas de hPRDX5. Seule la concentration en protéine a été modifiée (0.38 μM de hPRDX1/hPRDX2). La concentration en H_2O_2 a été gardée à 25 μM pour pouvoir faire une étude comparative avec hPRDX5 (même conditions). Nous avons par la suite vérifié que cela n'avait aucun impact significatif sur la mesure (Annexe 7).

L' $I'_{C_{50}}$ relatif du 4NPC a ensuite été déterminé pour ces deux protéines et les expériences ont menées à des $I'_{C_{50}}$ de $20 \pm 7 \mu\text{M}$ avec hPRDX1 et $25 \pm 10 \mu\text{M}$ avec hPRDX2 (Figure 94).

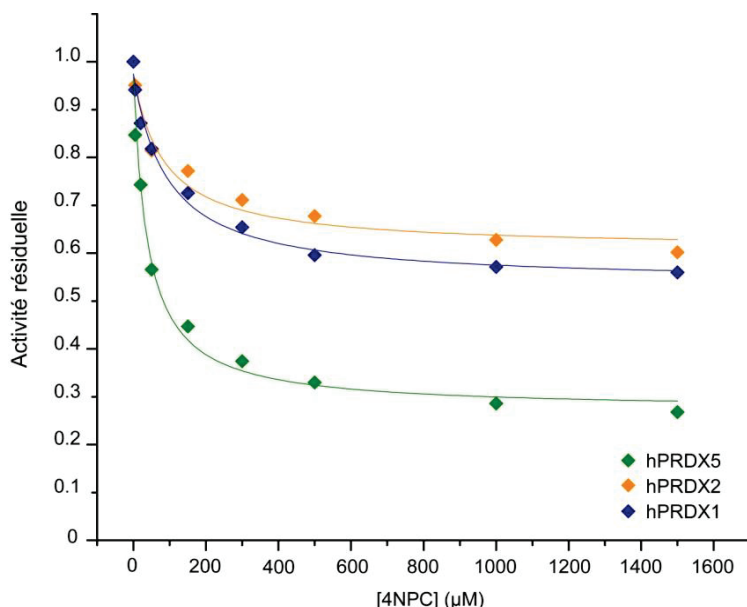


Figure 94. Comparaison de l'inhibition du ligand 4NPC pour hPRDX1, hPRDX2 et hPRDX5.

Courbes de détermination de $I'_{C_{50}}$ obtenues dans le cas du ligand 4NPC pour hPRDX1, hPRDX2 et hPRDX5. Vitesse de réaction en $\mu\text{M}/\text{min}/\text{L}$ normalisée pour la visualisation de l'activité résiduelle de chaque enzyme, exprimée en fonction de la concentration en ligand 4NPC introduite.

Les IC_{50} sont donc du même ordre de grandeur que ceux obtenus dans le cas de l'inhibition de hPRDX5, mais dans le cas de hPRDX1 et hPRDX2 on remarque que le pourcentage d'activité résiduelle (50 %) est supérieur à celui obtenu dans le cas de hPRDX5 (25 %). Dans le cas de hPRDX1 et hPRDX2, l'activité de l'enzyme est inhibée seulement de moitié par le 4NPC à saturation en ligand, alors que dans le cas de hPRDX5, 70 % de l'activité est inhibée. Si la mesure d' IC_{50} relatif ne nous permet pas de conclure quant à la sélectivité de l'inhibition, cette différence d'activité résiduelle indique une différence entre les isoformes hPRDX1 et hPRDX2 et l'isoforme hPRDX5, et donc une certaine sélectivité apparente.

- Sélectivité des autres ligands et comparaison

Nous avons mis en évidence une différence d'activité résiduelle des isoformes hPRDX1, hPRDX2 et hPRDX5 lors de leur inhibition par le ligand 4NPC, mais retrouve-t-on ce caractère sélectif au niveau des autres dérivés catéchols étudiés ? Pour répondre à cette question, les affinités des ligands CAT, 4MEC et 4TBC ont été estimées pour hPRDX1 et hPRDX2 par RMN. Etant des protéines oligomériques (de haut poids moléculaire), des expériences STD ont été réalisées, menant aux résultats présentés dans le Tableau 14 ci-dessous et comparés à ceux obtenus par ^{15}N -HSQC dans le cas de hPRDX5.

Globalement, les affinités déterminées par STD pour hPRDX1 et hPRDX2 sont beaucoup plus faibles que celles déterminées pour hPRDX5, atteignant presque la centaine de millimolaire pour l'interaction du ligand CAT avec hPRDX1. Si l'on compare les affinités obtenues pour hPRDX1 et hPRDX2, les résultats sont du même ordre de grandeur, avec une différence plus importante entre CAT et 4MEC/4TBC dans le cas de hPRDX1, et entre CAT/4MEC et 4TBC pour hPRDX2. Si l'on compare maintenant les K_D obtenus par STD et ^{15}N -HSQC, on remarque tout d'abord que le classement est similaire, le CAT étant le ligand le moins affiné et 4NPC le ligand le plus affiné. Les affinités déterminées, si elles sont du même ordre de grandeur, ne sont pas équivalentes à l'erreur près. Les différences observées peuvent venir de facteurs expérimentaux, tel que le seuil arbitraire utilisé en ^{15}N -HSQC, ou bien les protons utilisés pour le calcul du K_D en STD (les protons aromatiques uniquement). De plus, l'affinité par STD étant calculée par rapport au ligand, alors qu'elle est calculée par rapport aux acides aminés de la protéine en ^{15}N -HSQC, nous pouvons estimer que les affinités obtenues restent concordantes.

Vu les faibles affinités déterminées, uniquement l'inhibition du dérivé 4TBC a ensuite été étudiée pour ces deux systèmes protéiques. Des mesures d' IC_{50} ont été réalisées dans les mêmes conditions expérimentales que pour le 4NPC, et les résultats obtenus sont regroupés dans le Tableau 15 ci-dessous.

Tableau 14. Affinité des dérivés catéchols pour différentes protéines hPRDXs.

Affinités estimées par RMN ^1H STD des ligands CAT, 4MEC et 4TBC pour les protéines hPRDX1 et hPRDX2, et comparaison avec les valeurs obtenues pour hPRDX5.

Ligand	K_D hPRDX1 STD (mM)	K_D hPRDX2 STD (mM)	K_D hPRDX5 ^{15}N -HSQC (mM)	K_D hPRDX5 STD (mM)
CAT	$96 \pm 15^{\S}$	$56 \pm 10^{\S}$	$4.5 \pm 0.6^*$	5 ± 1
4MEC	$27 \pm 20^{\S}$	$54 \pm 13^{\S}$	$1.0 \pm 0.2^*$	0.5 ± 0.1
4TBC	$26 \pm 7^{\S}$	$5.9 \pm 1.0^{\S}$	$0.19 \pm 0.03^{\text{f}}$	0.37 ± 0.12

* Troussicot et al. *Journal of the American Chemical Society*. 2015¹⁸³

^{\S} Résultats non publiés.¹

^{\text{f}} M. Chow, L. Troussicot et al. *Biochemistry*. 2016.¹⁸⁸

Tableau 15. IC_{50} des dérivés 4TBC et 4NPC pour les isoformes hPRDX1, hPRDX2 et hPRDX5.

Valeurs d' IC_{50} (μM) et pourcentage d'activité résiduelle, obtenus dans le cas des ligands 4TBC et 4NPC avec hPRDX1 et hPRDX2, et comparaison avec les valeurs obtenues pour hPRDX5

Ligand		hPRDX1	hPRDX2	hPRDX5
4TBC ^{\text{f}}	IC_{50} relatif (μM)	240 ± 30	200 ± 30	430 ± 120
	Activité résiduelle	60 %	55 %	35 %
4NPC	IC_{50} relatif (μM)	20 ± 7	25 ± 10	40 ± 5
	Activité résiduelle	60 %	55 %	25 %

^{\text{f}} Chow M.L., Troussicot L. et al. 2016

Les IC_{50} relatifs obtenus pour l'inhibition des protéines hPRDX1 et hPRDX2 par le dérivé 4TBC sont meilleurs que celui déterminé pour l'inhibition de hPRDX5. En revanche, l'activité résiduelle est bien plus importante dans le cas de hPRDX1 et hPRDX2 qui ne sont inhibées que de moitié lorsque l'on atteint la saturation en inhibiteur. Cela pourrait suggérer que hPRDX5 est inhibée de manière sélective par les dérivés du catéchol 4TBC et 4NPC.

Nous avons donc vu grâce aux simulations FM et aux résultats obtenus par cinétique enzymatique que les dérivés catéchols avaient une affinité et une inhibition plus forte pour hPRDX5. Nous allons discuter l'origine structurale possible de ces différences observées.

Dans le chapitre I (III.1.1 p.53), nous avons décrit que les protéines de la famille des hPRDXs peuvent être classées selon le nombre de cystéines présentes dans leur structure, mais aussi selon le type de pont disulfure formé par les deux cystéines Cp et Cr des classes de hPRDXs à 2 cystéines.

hPRDX1 et hPRDX2 appartiennent à la classe des hPRDX à 2 cystéines typiques qui peuvent former des ponts disulfures intermoléculaires, tandis que hPRDX5 forme à elle seule la classe des hPRDX atypiques à 2 cystéines, formant des ponts disulfures intramoléculaires avec les cystéines présentes dans sa structure. De plus ces deux protéines ont une interface dimérique différente. hPRDX1 et hPRDX2 possèdent une interface dit de type B, par mise en contact de leurs feuillettes β tandis que hPRDX5 possède une interface dit de type A, par contact entre 2 hélices α (Figure 24 p.52). Cette différence peut jouer au niveau de la géométrie du site actif. En effet, si la cavité dans laquelle est localisée la Cp de hPRDX5 est plus étroite dû à la présence de l'hélice α formant l'interface, elle est aussi plus profonde, ce qui permet de maintenir la forme complexée par le biais d'interactions faibles de type vdW. L'hélice Leu116 – Phe120, seulement présente dans hPRDX5, permet de former des interactions de vdW supplémentaires avec le ligand et donc éventuellement d'augmenter son affinité. Cette hélice est absente dans le cas de hPRDX1 et hPRDX2. Au contraire, le site actif de hPRDX2 est beaucoup plus accessible, et situé plus en surface de la protéine, ce qui rend la cavité peu hydrophobe, et l'interaction affaiblie par les molécules de solvant qui diffusent au niveau de la cystéine catalytique et du ligand. Pour le ligand 4NPC par exemple, qui a été optimisé pour le site actif de hPRDX5 en créant des interactions avec l'hélice Leu116 – Phe120, sa sélectivité est d'autant plus grande que cette hélice n'est pas présente dans le cas des autres hPRDX. D'où son affinité moyenne pour hPRDX1 et hPRDX2, au même niveau que les ligands CAT ou 4TBC, mais fortement améliorée dans le cas de hPRDX5 par rapport à ces mêmes ligands.

- Estimation de la sélectivité du 4NPC par FM

Tout au long de ce projet, nous avons mis en évidence la corrélation des résultats obtenus par FM avec les résultats expérimentaux. Nous avons montré que cette technique permettait d'estimer non seulement les affinités PL de manière précise, mais aussi le type d'inhibition mis en jeu par le biais de l'étude des affinités pour l'enzyme libre et le complexe *ES*. En se basant toujours sur la détermination des affinités PL nous avons ensuite voulu utiliser la FM pour tenter de prédire la sélectivité de ces dérivés catéchols. Pour cela, nous avons cherché à estimer l'affinité de ces dérivés pour une autre isoforme. A l'aide de la FM, nous avons alors tenté d'estimer la sélectivité du ligand 4NPC en simulant son interaction avec une autre protéine de la famille des peroxyrédoxines.

D'après les structures disponibles dans la PDB pour les protéines hPRDX1 et hPRDX2, nous avons choisi de modéliser l'interaction du 4NPC avec hPRDX2. En effet, une structure de hPRDX2 sous forme réduite avec une résolution de 1.7 Å était disponible (1QMV.pdb), tandis que pour hPRDX1, les

structures disponibles étaient des protéines de mammifères, majoritairement sous forme oxydée, avec une moins bonne résolution. La simulation a ensuite été effectuée sur un dimère construit à partir de la structure du décimère comme décrit en chapitre II (p.79).

La simulation a été poursuivie jusqu'à atteindre la convergence de l'énergie, pour 600 ns. Durant cette simulation, 3 courts évènements de liaison seulement ont pu être observés. On s'attend donc à avoir une affinité bien moins bonne pour cette protéine. L'analyse de la trajectoire donne lieu à une valeur finale d'énergie libre de liaison de $\Delta G_b^0 = -3.0 \pm 0.5 \text{ kcal.mol}^{-1}$ soit un $K_D = 7.8 \pm 4.7 \text{ mM}$. On observe donc une diminution de l'affinité d'un facteur 100 au minimum par rapport à l'affinité déterminée pour hPRDX5, correspondant à une perte d'énergie libre de liaison d'environ 3 kcal.mol^{-1} visible sur la Figure 95.

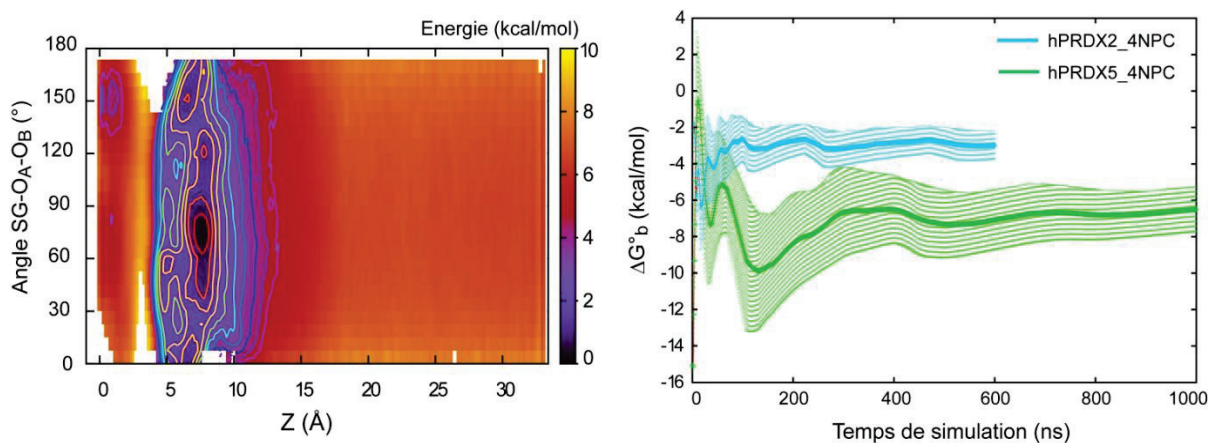


Figure 95. Analyse de l'énergie du complexe hPRDX2-4NPC lors de la simulation FM.

Gauche : Cartographie de l'énergie en fonction de la distance sur l'axe Z de l'entonnoir (CV1), et de l'angle $SG-O_A-O_B$ (CV3). Courbes de niveau représentées avec un pas de $0.5 \text{ kcal.mol}^{-1}$. Droite : Visualisation de la moyenne pondérée en fonction du temps de simulation de ΔG_b^0 pour le complexe hPRDX2-4NPC et comparaison avec le complexe hPRDX5-4NPC.

Si l'on cartographie cette énergie en fonction des variables collectives CV1 et CV3, on obtient de nombreux petits bassins d'énergie. Sont visibles un minimum énergétique à 8 \AA sur Z avec un angle d'environ 75° , et des minima d'énergie plus faible localisés au niveau du mur de potentiel avec un angle de 150° environ, et d'autres à 5 \AA avec des valeurs d'angles variables (Figure 95).

Par regroupement des différentes conformations de liaison dans le logiciel VMD, on observe les différentes interactions localisées dans ces minima (Figure 96). A 0 \AA sur Z, le ligand est en liaison avec Gly93 (cluster de type A). C'est une interaction secondaire qui a été observée durant cette simulation à hauteur de 29% des conformations liées. A 8 \AA sur Z, le ligand est en interaction avec Ser189 (chaîne B) ou Glu121 selon son orientation, pour des valeurs d'angle de 40° , 90° (minimum énergétique) et 120° environ (clusters de type B et D). L'acide aminé Ser189 appartient à l'hélice en *N-term* de la chaîne B qui est très mobile et qui vient recouvrir partiellement le site actif de la chaîne

A (Figure 96). Cette partie de la protéine n'a pas été contrainte lors de la modélisation, ce qui nous permet de mettre en évidence son rôle dans l'interaction PL puisque cette conformation compte pour 21% des conformations liées observées. L'interaction avec Glu121 qui est la plus fréquemment observée (30%) est comparable à l'interaction mise en évidence avec Glu179(B) dans le cas de hPRDX5. A 5 Å sur Z, le ligand est en interaction avec Cym50, pour des valeurs d'angle $SG-O_A-O_B$ variant entre 140° et 160°, selon la conformation mettant en jeu une seule liaison H (cluster de type C). La conformation d'angle $SG-O_A-O_B$ de 60° (liaison à Cym50 par deux liaisons H) est minoritaire, d'où le faible bassin énergétique correspondant sur la carte FES. La liaison avec Cys50 est dans le cas présent peu observée (seulement 20% des conformations liées), d'où l'affinité plus faible observée du 4NPC pour hPRDX2.

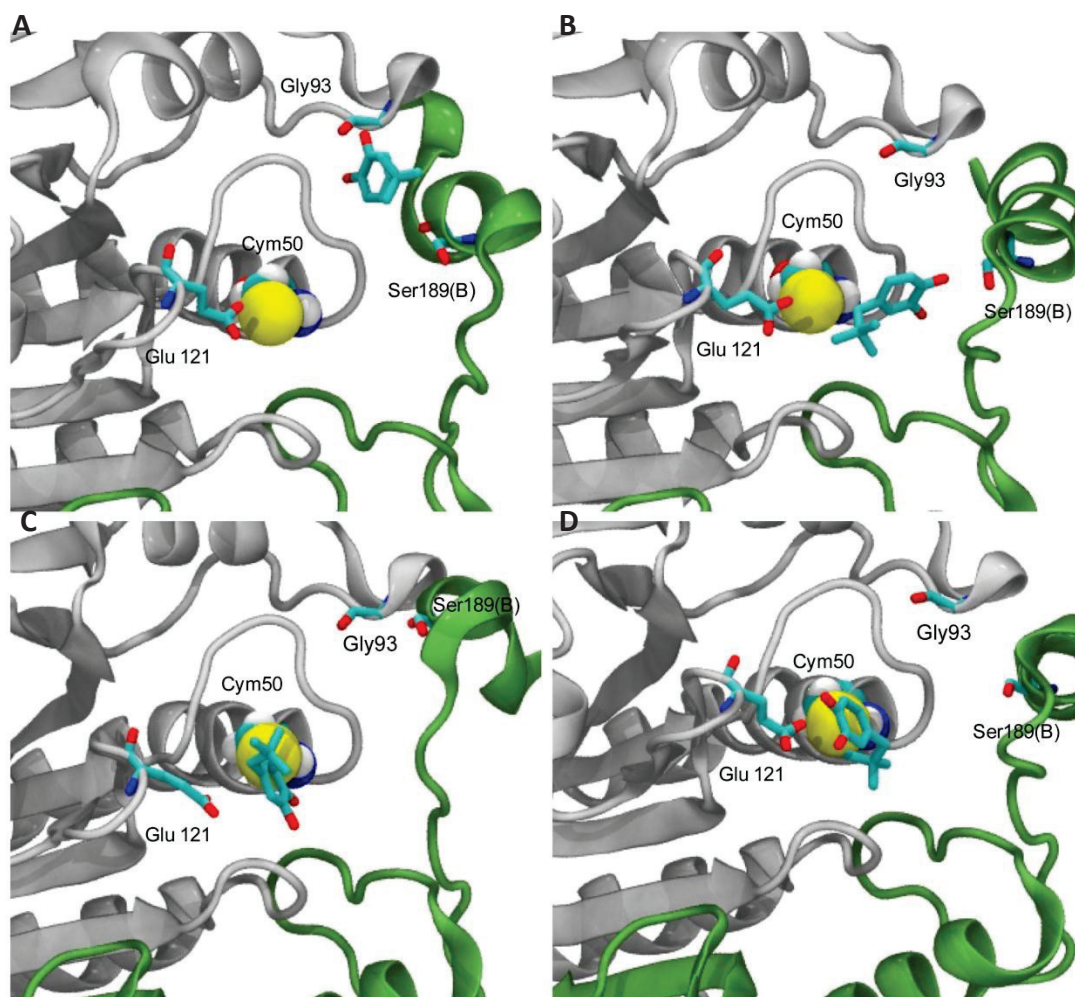


Figure 96. Visualisation des 4 conformations de liaison les plus fréquemment observées au cours de la simulation FM du complexe hPRDX2-4NPC.

Protéine colorée par chaîne monomérique : chaîne A = gris et chaîne B = vert. Atome SG Cym50 représenté sous forme d'une sphère de van der Waals jaune. Ligand 4NPC et résidus impliqués dans l'interaction colorés selon la nature de leurs atomes (C = cyan ; O = rouge ; N = bleu). Proportion des différents clusters : A = 29% ; B = 21% ; C = 20% ; D = 30%.

Ayant préalablement déterminé l'inhibition de hPRDX2 par le dérivé 4NPC, nous pouvons corréler ces résultats avec l'affinité apparente estimée de 4NPC pour hPRDX2 qui est nettement plus faible que celle estimée pour hPRDX5. En effet, nous avons vu que ce dérivé inhibait ces deux isoformes avec un IC_{50} relatif similaire, mais une activité résiduelle deux fois plus importante dans le cas de hPRDX2 qui est inhibée à hauteur de 50% de son activité au maximum, tandis que hPRDX5 peut voir son activité inhibée jusqu'à 75%. La différence d'inhibition induite par le 4NPC est donc bien le reflet de la différence d'affinité observée en FM.

CHAPITRE IV

CONCLUSION ET PERSPECTIVES

I. CONCLUSION

L'objectif général de cette étude était d'identifier de nouveaux inhibiteurs potentiels pour le traitement des chocs ischémiques, et plus précisément contre les hPRDX proposées comme étant impliquées dans l'inflammation post-ischémique cérébrale. Les études précédentes réalisées au laboratoire ayant conduit à l'identification du catéchol comme se liant au site actif de l'une des protéines de cette famille : hPRDX5, ma recherche a été centrée sur l'interaction de cette protéine dimérique, avec des petites molécules dérivées du catéchol. Dans ce concept d'identification et d'optimisation de petites molécules, nous avons utilisé un protocole de métadynamique, la FM, afin de guider notre recherche et de montrer son potentiel en chimie médicinale pour la conception de molécules bioactives.

Dans un premier temps, nous avons montré que cette technique de dynamique moléculaire permettait l'estimation précise des affinités PL, dont les résultats corrèlent avec ceux obtenus expérimentalement par RMN en nous appuyant sur l'étude de deux molécules ligands connues.

Ces résultats nous ont permis dans un deuxième temps de pouvoir nous appuyer sur les informations disponibles grâce à cette technique telles que les différentes conformations de liaison, la cartographie de l'énergie libre de liaison, l'expansion du site actif ainsi que la statistique de liaison, afin de guider l'optimisation de ces molécules fragments dérivées du catéchol. De plus, la visualisation des trajectoires du ligand à l'intérieur du site actif nous ont permis de mettre en évidence des conformations de liaison en échange permanent ainsi que des interactions secondaires permettant la stabilisation du complexe PL. L'utilisation de la FM pour la conception de molécules affines s'est révélée efficace puisque nous avons réussi à optimiser le fragment catéchol jusqu'à atteindre des affinités de l'ordre du micromolaire en stabilisant la forme liée du complexe, et en améliorant les interactions hydrophobes présentes à l'intérieur de la cavité afin de favoriser une conformation de liaison particulière. Ces résultats ont été confirmés par RMN ¹⁵N-HSQC et par ¹H-STD dans le cas du 4TBC et du 4NPC notamment.

Par la suite, l'affinité n'étant pas un critère suffisant pour prouver et mesurer l'inhibition d'une enzyme, nous avons étudié la force inhibitrice de ces dérivés pour hPRDX5 et deux autres isoformes,

hPRDX1 et hPRDX2. Pour aller plus loin dans l'utilisation de la FM, nous avons mis en place des protocoles afin d'estimer l'inhibition des ligands étudiés par le biais du calcul. Bien que les observations de dynamique ne puissent être directement reliées à la force de l'inhibition, elles peuvent être utilisées afin de prédire le mécanisme de cette inhibition. Les expériences de FM effectuées dans le cadre de mon étude ont montrées que les molécules testées avaient la possibilité de se lier à l'enzyme libre comme au complexe de Michaelis *ES*, laissant à penser que les dérivés catéchols inhibaient hPRDX5 avec un mécanisme non compétitif. Ces résultats prédictifs ont pu être corrélés avec des résultats expérimentaux obtenus par cinétique enzymatique qui nous ont permis de déterminer la force de l'inhibition des ces dérivés avec pour hPRDX1, hPRDX2 et hPRDX5 via la mesure de leur IC_{50} et la détermination de leur mécanisme d'inhibition pour hPRDX5 partiel mixte non compétitif.

Les différences observées dans l'affinité et l'inhibition des protéines hPRDX1, hPRDX2 et hPRDX5 peuvent être attribuées aux différences structurales de ces protéines. En effet, nous avons vu que hPRDX1 et hPRDX2 issues de la même classe de hPRDX étaient inhibées avec une activité similaire, différente de celle observée pour hPRDX5. Une sélectivité des dérivés catéchols pour hPRDX5 a donc pu être observée sur la base de ces trois systèmes, mais qu'en est-il des autres protéines hPRDX ? De plus, nous avons étudié l'affinité du substrat naturel de hPRDX5 : H_2O_2 , et l'inhibition des dérivés catéchols lors de la réaction enzymatique avec ce substrat. Nous pouvons cependant nous demander si l'utilisation d'un substrat différent pourrait amener des modifications dans le mécanisme d'inhibition de ces protéines.

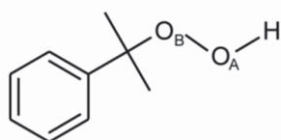
II. PERSPECTIVES

Les résultats obtenus sur l'étude de l'affinité et de l'inhibition de hPRDX5 par les dérivés catéchols suscitent des nouvelles perspectives et questions concernant les isoformes non abordées ici de cette famille de protéines (hPRDX3, 4 et 6), l'étude de substrats différents dans la réaction enzymatique catalysée, et l'optimisation des dérivés catéchols par la recherche de motifs structuraux spécifiques ou l'étude de molécules naturelles.

- Etude d'un substrat différent : le peroxyde de cumène

Toutes les expériences menées sur l'inhibition des isoformes hPRDX1 hPRDX2 et hPRDX5 ont été effectuées avec le peroxyde d'hydrogène H_2O_2 , substrat de hPRDX5. Cependant nous savons qu'il existe d'autres peroxydes organiques tel que l'hydroperoxyde de cumène par exemple, qui peuvent être substrat de la réaction enzymatique catalysée par les PRDX. Il serait donc intéressant de voir si l'utilisation d'un tel substrat amène des différences significatives dans l'inhibition de ces protéines.

Le peroxyde de cumène (formule ci-dessous) est un autre substrat connu des PRDXs bactériennes (AhpC) et humaines (hPRDX1, hPRDX2 et hPRDX6).^{112; 125} De manière préliminaire, nous avons modélisé son interaction et estimé son affinité avec hPRDX5. De plus, son étude avec hPRDX5 pourrait mener à des informations sur la sélectivité des substrats des hPRDX, par rapport à H_2O_2 dans le cas de l'isoforme hPRDX5



Hydroperoxyde de cumène

Une simulation de 300 ns de l'hydroperoxyde de cumène avec hPRDX5 a été effectuée et n'a donné lieu qu'à une seule période de liaison. Comme précédemment, nous avons utilisé trois variables collectives pour définir le système : CV1 représente la distance sur l'axe Z, CV2 la distance par rapport à cet axe et CV3 la valeur de l'angle SG-O_A-O_B.

La FES calculée en fonction des variables CV1 et CV3, un seul minimum énergétique est visible à 5 Å sur l'axe Z, représentant l'interaction du peroxyde de cumène avec Cys47, avec un angle de 110° (Figure 97). En effet, l'utilisation d'un peroxyde ne rend pas possible la conformation de liaison à

deux liaisons H observée dans le cas des dérivés du catéchol. L'affinité du cumène pour hPRDX5 calculée par cette simulation est de $\Delta G_b^\circ = -2.1 \pm 0.1 \text{ kcal.mol}^{-1}$ soit une constante de dissociation du complexe, $K_D = 29 \pm 3 \text{ mM}$. Comme pour beaucoup d'enzymes, l'affinité pour le substrat reste faible (cas du peroxyde de cumène) ou voisine de zéro (cas de H_2O_2). On remarque que l'affinité du peroxyde de cumène pour hPRDX5 est plus grande que celle de H_2O_2 . Nous pouvons nous demander si cette différence pourrait induire une modification du mécanisme d'inhibition des hPRDX par les dérivés catéchols si le peroxyde de cumène était utilisé comme substrat.

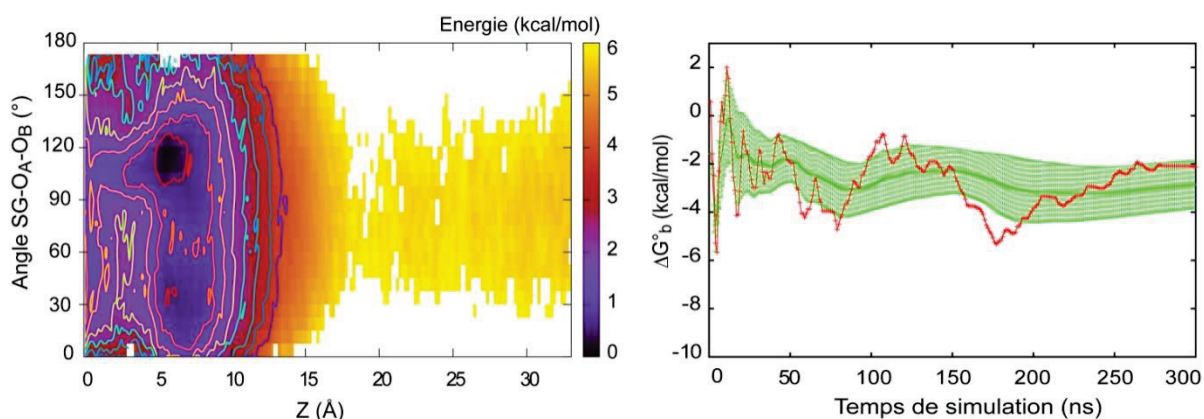


Figure 97. Analyse de l'énergie du complexe hPRDX5-peroxyde de cumène par FM.

Gauche : Cartographie de l'énergie en fonction de la distance sur l'axe Z de l'entonnoir (CV1), et de l'angle SG-OA-OB (CV3). Courbes de niveau représentées avec un intervalle de $0.5 \text{ kcal.mol}^{-1}$. Droite : Visualisation de la moyenne pondérée en fonction du temps de simulation de ΔG_b° pour le complexe hPRDX5-peroxyde de cumène.

A la suite de ces premières estimations par FM, des tests préliminaires expérimentaux ont été effectués en cinétique enzymatique. La courbe expérimentale obtenue ne correspond pas au modèle hyperbolique de Michaelis-Menten (Figure 98) mais plutôt à un modèle sigmoïdal, que nous avons fait correspondre avec une courbe théorique de type :

$$y = V_{min} + (V_{max} - V_{min}) \frac{x^n}{k_M^n + x^n} \quad [42]$$

Par correspondance avec cette équation, on trouve un K_M de l'ordre de $3 \mu\text{M}$, pour un V_{max} de $33 \mu\text{mol.min}^{-1}.\text{L}^{-1}$. Le K_M est donc plus faible que celui mesuré dans le cas de H_2O_2 .

Ces tests nous indiquent que le cumène induirait un caractère coopératif (coopérativité positive) de hPRDX5 au contraire de H_2O_2 ce qui met en évidence une limitation. En effet, nous prenons uniquement en compte un site de fixation dans nos expériences, considérant un modèle de liaison 1:1. Ce modèle est bien approprié pour H_2O_2 , mais ne le sera pas dans le cas du peroxyde de cumène.

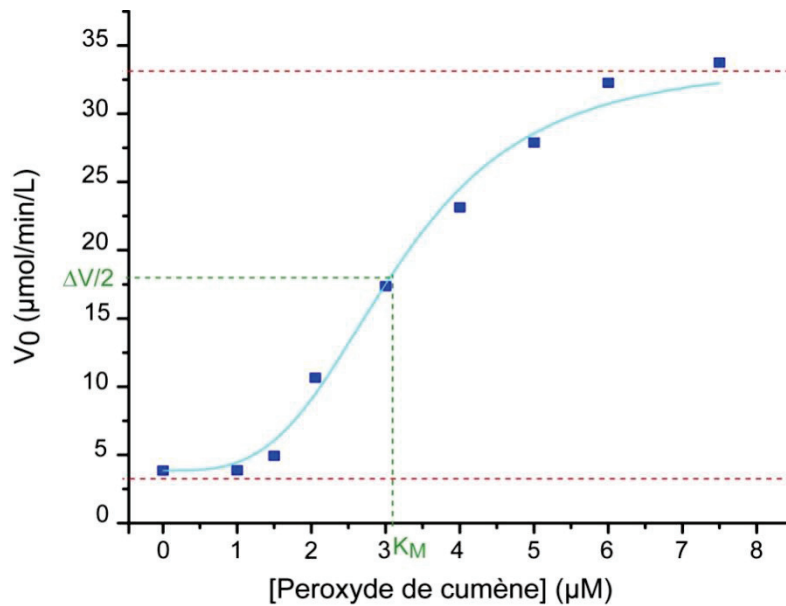


Figure 98. Courbe de cinétique effectuée avec le peroxyde de cumène comme substrat de la réaction enzymatique catalysée par hPRDX5.

$V_0 = f([\text{peroxyde de cumène}])$. Vitesse initiale de la réaction en $\mu\text{m}\cdot\text{min}^{-1}\cdot\text{L}^{-1}$. Correspondance avec une courbe théorique de type : $y = V_{min} + (V_{max} - V_{min}) x^n / K_M^n + x^n$ pour la détermination de K_M .

D'après la courbe obtenue de manière préliminaire, on peut calculer l'efficacité enzyme qui est de $k_{cat}/K_M = 8.10^5 \text{ M}^{-1}\cdot\text{s}^{-1}$, presque 10 fois plus grande que celle obtenue avec H_2O_2 $k_{cat}/K_M = 9.10^4 \text{ M}^{-1}\cdot\text{s}^{-1}$.¹⁸⁸ Cela est cohérent avec l'étude menée par Parsonage *et al.*¹²⁵ ainsi qu'avec les résultats publiés par Trujilo *et al.* sur la cinétique de réaction de hPRDX5 avec les peroxydes organiques par rapport à H_2O_2 .¹⁹⁰

Il serait tout de même nécessaire de refaire cette expérience afin de confirmer ou non la coopérativité du cumène avec hPRDX5, et ensuite, de voir si ce caractère diffère dans l'étude de ce même substrat avec d'autres isorformes de hPRDXs.

- Réflexion sur la relation temps de résidence/affinité

Dans ce projet, nous avons principalement exploré le caractère thermodynamique du phénomène d'association-dissociation protéine-ligand par FM en estimant l'énergie libre du système au cours du temps. Nous savons que l'affinité n'est pas uniquement un bilan de la thermodynamique de l'équilibre d'interaction, mais aussi de sa cinétique. Pour cela on cherche généralement à estimer la valeur des constantes de vitesse k_{on} et k_{off} de cet équilibre. Estimer ces constantes de vitesses par le biais de la MD/FM est une motivation de nombreux projets de recherche. Des travaux publiés récemment par Tiwary *et al.* ont montré qu'il était possible d'obtenir une estimation de ces

constantes.⁵⁰ Pour obtenir une estimation du k_{on} il faut partir du complexe libre, et continuer la simulation jusqu'à observer l'association du complexe. Au contraire, pour avoir une estimation du k_{off} il faut partir du complexe lié, et observer sa dissociation. La moyenne des temps d'association (ou de dissociation) observés sur l'ensemble des simulations effectuées (environ 20 dans leur cas) permettrait une estimation précise des constantes de vitesse, dont la distribution répondrait au test de Kolmogorov-Smirnov.⁵⁰

Par le biais de la cinétique de l'équilibre, un autre paramètre peut être pris en compte : le temps de résidence t_R . Nous avons vu dans le chapitre I que cette grandeur pouvait être reliée au k_{off} (relation 9 p.10), et donc au K_D . Nous savons qu'un inhibiteur n'est actif que lorsqu'il est lié à sa cible. Donc l'augmentation du \bar{t}_R , correspond à une diminution du k_{off} si l'on considère que la cinétique de l'interaction est limitée par la diffusion ($k_{on} \approx 10^{10} \text{ M}^{-1} \cdot \text{s}^{-1}$), et sera synonyme d'une meilleure activité.

En FM, le \bar{t}_R estimé sur les événements de liaison observés n'est que purement qualitatif. Ces événements d'association-dissociation sont aléatoires. Evidemment plus ils prennent de temps et moins ils sont nombreux dans un laps de temps donné. Nous avons supposé que le temps de résidence se distribue selon une Loi Normale. Les potentiels de biais appliqués en métadynamique vont biaiser le \bar{t}_R observé et tendre à diminuer le k_{off} puisque l'énergie est déposée afin de dissocier le complexe. Ainsi les événements observés ne peuvent nous permettre de calculer un \bar{t}_R réaliste et quantitatif, mais peuvent permettre une comparaison qualitative des \bar{t}_R apparents observés en FM puisque ces simulations ont été effectuées avec les mêmes paramètres de biais. Les \bar{t}_R estimés par FM sur les simulations effectuées, répertoriés dans le Tableau 16, montrent une augmentation générale du temps de résidence conjointe à l'augmentation de l'affinité des ligands correspondants.

Tableau 16. Evolution du \bar{t}_R mesuré par FM en fonction de l'affinité estimée pour hPRDX5, pour les principaux dérivés catéchols étudiés

Ligand	Nombre d'évènements	\bar{t}_R (ns)	ΔG°_b estimée (kcal/mol)	K_D estimé (M)	k_{off} (s^{-1})
CAT	6	11	- 3.0	$6.9 \cdot 10^{-3}$	$6.9 \cdot 10^8$
4MEC	12	13	- 3.6	$2.4 \cdot 10^{-3}$	$2.4 \cdot 10^8$
4TBC	6	22	-5.5	$1.8 \cdot 10^{-4}$	$1.8 \cdot 10^7$
4NPC	6	43	- 6.6	$2.0 \cdot 10^{-5}$	$2.0 \cdot 10^6$
4N5PC	5	100	-8.2	$2.0 \cdot 10^{-6}$	$2.0 \cdot 10^5$
Hit 3	6	41	-8.2	$3.0 \cdot 10^{-6}$	$3.0 \cdot 10^5$

Avec $k_{off} = K_D \times k_{on}$ et k_{on} estimé à $10^{10} \text{ M}^{-1} \cdot \text{s}^{-1}$

L'exception du ligand Hit 3, dont le \bar{t}_R est plus faible que celui calculé pour 4N5PC pour une affinité calculée équivalente, illustre le côté aléatoire de la dynamique, et la distribution normale du temps de résidence. Pour avoir une analyse plus précise des \bar{t}_R il faudrait suivre le protocole utilisé par Tiwary et al. en effectuant un nombre suffisant de simulations nécessaires à l'obtention d'une moyenne plus réaliste.⁵⁰

- Etude des différentes isoformes de hPRDX

Nous avons proposé une estimation de la sélectivité des ligands catéchols par rapport à deux autres isoformes des hPRDX : hPRDX1 et hPRDX2. Nous avons vu qu'en terme d'inhibition les IC_{50} déterminés étaient similaires à ceux trouvés pour hPRDX5, mais que l'activité résiduelle était plus importante pour ces deux isoformes de la classe des hPRDX à 2-cystéines typiques. Cependant des expériences complémentaires permettraient de lever certaines ambiguïtés concernant l'affinité et l'inhibition des autres isoformes. Cela passe par une meilleure compréhension de l'oligomérisation de ces protéines, ou encore la mise en place de dynamique moléculaire non biaisée qui pourrait mettre en évidence d'autres régions (autre que le site récepteur) dans lesquelles les dérivés catéchols pourraient se lier.

De plus, si nous avons estimé la sélectivité pour les isoformes hPRDX1 et hPRDX2, aucune étude n'a été menée concernant les isoformes hPRDX3, hPRDX4 (même classe que hPRDX1 et hPRDX2) et hPRDX6 (hPRDX à 1 cystéine). hPRDX6 étant une protéine monomérique, il serait théoriquement aisé de modéliser son interaction avec les dérivés catéchols à partir d'une structure cristallographique résolue. Une structure de cette isoforme dans son état réduit a d'ailleurs récemment été déposée à la PDB avec une résolution de 2.5 Å (5B6M.pdb).¹⁹¹ Cette isoforme monomérique ne possède pas d'interface comme dans les autres isoformes humaines, et il serait intéressant de voir et de comprendre l'impact de l'oligomérisation de ces enzymes sur l'interaction grâce à la dynamique moléculaire.

- Inhibiteurs de hPRDX dérivés du catéchol

Partant du fragment catéchol (110 g.mol⁻¹) ayant une affinité de l'ordre du millimolaire, nous avons développé de manière minimale ce fragment pour atteindre une affinité de l'ordre du micromolaire avec le 4NPC. Dans un contexte de conception de médicaments, ce fragment doit ensuite être développé pour obtenir une molécule de type *drug-like* qui aurait une affinité pouvant atteindre le nanomolaire pour être candidat médicament. Dans ce projet, nous avons développé et optimisé le fragment 4NPC et obtenu des estimations encourageantes d'affinités obtenues par FM, qui n'ont pas encore pu être corrélées à l'expérience. Des expériences de RMN et de cinétique

enzymatique sur ces dérivés de type *drug-like* 4N5PC et Hit 3 seraient donc bénéfiques pour la suite du développement d'inhibiteurs dérivés du catéchol.

De manière parallèle, une étude rapide de molécules naturelles possédant le motif catéchol pourrait être réalisée. En effet, nous avons vu en chapitre d'introduction que certains neurotransmetteurs et polyphénols naturels possèdent un motif catéchol (structures en Figure 35 p.81). Ainsi, de manière préliminaire, la dopamine et l'adrénaline ont été modélisées par MD non biaisée (NVT) avec hPRDX5. A l'issue de ces simulations, aucun évènement de liaison significatif n'a pu être observé avec les Cys47 des chaînes A et B de l'homodimère hPRDX5. Néanmoins, les molécules naturelles peuvent offrir de nombreuses possibilités de par leur variété, et donc il serait intéressant de tester d'autres molécules possédant un motif catéchol (voir chapitre I p.75). Cela pourrait permettre de mettre en évidence une corrélation entre des motifs structuraux et l'inhibition des hPRDX (SAR), et d'aller ainsi encore plus loin dans l'optimisation des dérivés catéchols étudiés.

BIBLIOGRAPHIE

1. Chow, M. L. (2016). Compréhension des effets de l'inhibitions des protéines peroxyrédoxines humaines pour le traitement potentiel de l'inflammation post-ischemique du cerveau, Université de Lyon.
2. Williams, M. A. & Daviter, T. (2013). *Protein-Ligand Interactions Methods and Applications*, Humana Press : Imprint: Humana Press, Totowa, NJ.
3. Nienhaus, G. U. (2005). *Protein-ligand interactions : methods and applications*, Humana Press, Totowa, N.J.
4. Kuriyan, K. W. (2012). Molecular recognition : the thermodynamics of binding. In *The Molecules of Life - Physical and Chemical Principles*. Garland Science.
5. Pan, A. C., Borhani, D. W., Dror, R. O. & Shaw, D. E. (2013). Molecular determinants of drug-receptor binding kinetics. *Drug Discovery Today* **18**, 667-673.
6. Guldberg, C. M., Waage, P. (1867). Etudes sur les affinités chimiques. *Christiana : Brogger & Χηριστιε*.
7. Ehrlich, P. (1913). Chemotherapeutics: Scientific Principles, Methods and Results. *Lancet Neurol* **182**, 445.
8. Alberty, R. A. & Hammes, G. G. (1958). Application of the Theory of Diffusion-controlled Reactions to Enzyme Kinetics. *The Journal of Physical Chemistry* **62**, 154-159.
9. Zhou, G.-Q. & Zhong, W.-Z. (1982). Diffusion-Controlled Reactions of Enzymes. *European Journal of Biochemistry* **128**, 383-387.
10. Eyring, H. (1935). The activated complex in chemical reaction. *J Chem Phys* **3**, 9.
11. Fisher, E. (1894). *Ber.* **27**, 2985.
12. Koshland, D. E. (1958). Application of a Theory of Enzyme Specificity to Protein Synthesis. *Proceedings of the National Academy of Sciences of the United States of America* **44**, 98-104.
13. Sullivan, S. M. & Holyoak, T. (2008). Enzymes with lid-gated active sites must operate by an induced fit mechanism instead of conformational selection. *Proceedings of the National Academy of Sciences* **105**, 13829-13834.
14. Karush, F. (1950). Heterogeneity of the Binding Sites of Bovine Serum Albumin1. *Journal of the American Chemical Society* **72**, 2705-2713.
15. Monod, J., Wyman, J. & Changeux, J. P. (1965). ON THE NATURE OF ALLOSTERIC TRANSITIONS: A PLAUSIBLE MODEL. *J Mol Biol* **12**, 88-118.
16. Cerny, J. & Hobza, P. (2007). Non-covalent interactions in biomacromolecules. *Phys Chem Chem Phys* **9**, 5291-303.
17. Frieden, E. (1975). Non-covalent interactions: Key to biological flexibility and specificity. *Journal of Chemical Education* **52**, 754.
18. Cooper, M. A. (2002). Optical biosensors in drug discovery. *Nat Rev Drug Discov* **1**, 515-28.

19. Freyer, M. W. & Lewis, E. A. (2008). Isothermal titration calorimetry: experimental design, data analysis, and probing macromolecule/ligand binding and kinetic interactions. *Methods Cell Biol* **84**, 79-113.
20. Pierce, M. M., Raman, C. S. & Nall, B. T. (1999). Isothermal titration calorimetry of protein-protein interactions. *Methods* **19**, 213-21.
21. Jelesarov, I. & Bosshard, H. R. (1999). Isothermal titration calorimetry and differential scanning calorimetry as complementary tools to investigate the energetics of biomolecular recognition. *Journal of Molecular Recognition* **12**, 3-18.
22. Calleri, E., Fracchiolla, G., Montanari, R., Pochetti, G., Lavecchia, A., Loiodice, F., Laghezza, A., Piemontese, L., Massolini, G. & Temporini, C. (2012). Frontal affinity chromatography with MS detection of the ligand binding domain of PPARgamma receptor: ligand affinity screening and stereoselective ligand-macromolecule interaction. *J Chromatogr A* **1232**, 84-92.
23. Vuignier, K., Schappler, J., Veuthey, J. L., Carrupt, P. A. & Martel, S. (2010). Drug-protein binding: a critical review of analytical tools. *Anal Bioanal Chem* **398**, 53-66.
24. Erlanson, D. A., McDowell, R. S. & O'Brien, T. (2004). Fragment-based drug discovery. *J Med Chem* **47**, 3463-82.
25. Yu, C., Chen, L., Luo, H., Chen, J., Cheng, F., Gui, C., Zhang, R., Shen, J., Chen, K., Jiang, H. & Shen, X. (2004). Binding analyses between Human PPARgamma-LBD and ligands. *Eur J Biochem* **271**, 386-97.
26. Caliendo, R., Belviso, D. B., Aresta, B. M., de Candia, M. & Altomare, C. D. (2013). Protein crystallography and fragment-based drug design. *Future Med Chem* **5**, 1121-40.
27. Nienaber, V. L., Richardson, P. L., Klighofer, V., Bouska, J. J., Giranda, V. L. & Greer, J. (2000). Discovering novel ligands for macromolecules using X-ray crystallographic screening. *Nat Biotechnol* **18**, 1105-8.
28. Wilson, W. D. (2002). Tech.Sight. Analyzing biomolecular interactions. *Science* **295**, 2103-5.
29. Pellecchia, M., Bertini, I., Cowburn, D., Dalvit, C., Giralt, E., Jahnke, W., James, T. L., Homans, S. W., Kessler, H., Luchinat, C., Meyer, B., Oschkinat, H., Peng, J., Schwalbe, H. & Siegal, G. (2008). Perspectives on NMR in drug discovery: a technique comes of age. *Nature reviews. Drug discovery* **7**, 738-745.
30. Meyer, B. & Peters, T. (2003). NMR spectroscopy techniques for screening and identifying ligand binding to protein receptors. *Angew Chem Int Ed Engl* **42**, 864-90.
31. Kay, L., Keifer, P. & Saarinen, T. (1992). Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *Journal of the American Chemical Society* **114**, 10663-10665.
32. Morris, G. A. & Freeman, R. (1979). Enhancement of nuclear magnetic resonance signals by polarization transfer. *Journal of the American Chemical Society* **101**, 760-762.
33. Mayer, M. M., B. (1999). <Mayer&Meyer_Angew.Chem.Int.Ed_1999.pdf>. *Angew Chem Int Ed* **38**, 5.
34. Mayer, M. & Meyer, B. (2001). Group epitope mapping by saturation transfer difference NMR to identify segments of a ligand in direct contact with a protein receptor. *J Am Chem Soc* **123**, 6108-17.
35. Chipot, C. (2009). Molecular Dynamics. Observing Matter in Motion. In *Nanoscience: Nanobiotechnology and Nanobiology* (Boisseau, P., Houdy, P. & Lahmani, M., eds.), pp. 803-838. Springer Berlin Heidelberg, Berlin, Heidelberg.
36. Leach, A. R. (2001). *Molecular modelling : principles and applications*, Prentice Hall, Harlow [etc.].

37. Oppenheimer, M. B. a. R. (1927). Zur Quantentheorie der Molekel. In *Annalen der Physik* (Barth, J. A., ed.), pp. 28. Gilbert, Wilhelm, Germany.
38. Chipot, C. (2002). Les méthodes numériques de la dynamique moléculaire, pp. 25.
39. Harvey, M. J. & De Fabritiis, G. (2009). An Implementation of the Smooth Particle Mesh Ewald Method on GPU Hardware. *J Chem Theory Comput* **5**, 2371-7.
40. Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W. & Kollman, P. A. (1995). A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *Journal of the American Chemical Society* **117**, 5179-5197.
41. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S. & Karplus, M. (1983). CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *Journal of Computational Chemistry* **4**, 187-217.
42. Durrant, J. D. & McCammon, J. A. (2011). Molecular dynamics simulations and drug discovery. *BMC Biol* **9**, 71.
43. Karplus, M. (2010). Dynamical aspects of molecular recognition. *Journal of Molecular Recognition* **23**, 102-104.
44. Andersen, H. C. (1980). Molecular Dynamics Simulations at Constant Pressure and/or Temperature. *J Chem Phys* **72**.
45. Deng, Y. & Roux, B. (2009). Computations of standard binding free energies with molecular dynamics simulations. *J Phys Chem B* **113**, 2234-46.
46. Buch, I., Sadiq, S. K. & De Fabritiis, G. (2011). Optimized Potential of Mean Force Calculations for Standard Binding Free Energies. *J Chem Theory Comput* **7**, 1765-72.
47. Gumbart, J. C., Roux, B. & Chipot, C. (2013). Efficient Determination of Protein–Protein Standard Binding Free Energies from First Principles. *Journal of Chemical Theory and Computation* **9**, 3789-3798.
48. Gumbart, J. C., Roux, B. & Chipot, C. (2013). Standard Binding Free Energies from Computer Simulations: What Is the Best Strategy? *Journal of Chemical Theory and Computation* **9**, 794-802.
49. Aqvist, J., Luzhkov, V. B. & Brandsdal, B. O. (2002). Ligand binding affinities from MD simulations. *Acc Chem Res* **35**, 358-65.
50. Sham, Y. Y., Chu, Z. T., Tao, H. & Warshel, A. (2000). Examining methods for calculations of binding free energies: LRA, LIE, PDL-D-LRA, and PDL-D/S-LRA calculations of ligands binding to an HIV protease. *Proteins* **39**, 393-407.
51. Decherchi, S., Masetti, M., Vyalov, I. & Rocchia, W. (2015). Implicit solvent methods for free energy estimation. *Eur J Med Chem* **91**, 27-42.
52. Rao, S. N., Singh, U. C., Bash, P. A. & Kollman, P. A. (1987). Free energy perturbation calculations on binding and catalysis after mutating Asn 155 in subtilisin. *Nature* **328**, 551-4.
53. Tiwary, P., Limongelli, V., Salvalaglio, M. & Parrinello, M. (2015). Kinetics of protein–ligand unbinding: Predicting pathways, rates, and rate-limiting steps. *Proceedings of the National Academy of Sciences* **112**, E386-E391.
54. Tiwary, P., Mondal, J., Morrone, J. A. & Berne, B. J. (2015). Role of water and steric constraints in the kinetics of cavity–ligand unbinding. *Proceedings of the National Academy of Sciences* **112**, 12015-12019.

55. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics* **79**, 926-935.
56. Nada, H. & van der Eerden, J. P. J. M. (2003). An intermolecular potential model for the simulation of ice and water near the melting point: A six-site model of H₂O. *The Journal of Chemical Physics* **118**, 7401-7413.
57. Shaw, D. E., Deneroff, M. M., Dror, R. O., Kuskin, J. S., Larson, R. H., Salmon, J. K., Young, C., Batson, B., Bowers, K. J., Chao, J. C., Eastwood, M. P., Gagliardo, J., Grossman, J. P., Ho, C. R., Ierardi, D. J., Istv, #225, Kolossv, n., #225, ry, Klepeis, J. L., Layman, T., McLeavey, C., Moraes, M. A., Mueller, R., Priest, E. C., Shan, Y., Spengler, J., Theobald, M., Towles, B. & Wang, S. C. (2008). Anton, a special-purpose machine for molecular dynamics simulation. *Commun. ACM* **51**, 91-97.
58. Giupponi, G., Harvey, M. J. & De Fabritiis, G. (2008). The impact of accelerator processors for high-throughput molecular modeling and simulation. *Drug Discov Today* **13**, 1052-8.
59. Harvey, M. J. & De Fabritiis, G. (2012). High-throughput molecular dynamics: the powerful new tool for drug discovery. *Drug Discovery Today* **17**, 1059-1062.
60. Harvey, M. J., Giupponi, G. & Fabritiis, G. D. (2009). ACEMD: Accelerating Biomolecular Dynamics in the Microsecond Time Scale. *J Chem Theory Comput* **5**, 1632-9.
61. Barducci, A., Bonomi, M. & Parrinello, M. (2011). Metadynamics. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **1**, 826-843.
62. Laio, A. & Parrinello, M. (2002). Escaping free-energy minima. *Proc Natl Acad Sci USA* **99**, 12562-6.
63. Laio, A. & Gervasio, F. L. (2008). Metadynamics: a method to simulate rare events and reconstruct the free energy in biophysics, chemistry and material science. *Reports on Progress in Physics* **71**, 126601.
64. Bonomi, M., Branduardi, D., Bussi, G., Camilloni, C., Provasi, D., Raiteri, P., Donadio, D., Marinelli, F., Pietrucci, F., Broglia, R. A. & Parrinello, M. (2009). PLUMED: A portable plugin for free-energy calculations with molecular dynamics. *Computer Physics Communications* **180**, 1961-1972.
65. Gervasio, F. L., Laio, A. & Parrinello, M. (2005). Flexible docking in solution using metadynamics. *J Am Chem Soc* **127**, 2600-7.
66. Johannes, K. (2011). Umbrella Sampling. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **00**, 11.
67. Kumar, S., Rosenberg, J. M., Bouzida, D., Swendsen, R. H. & Kollman, P. A. (1992). THE weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *Journal of Computational Chemistry* **13**, 1011-1021.
68. Izrailev, S., Crofts, A. R., Berry, E. A. & Schulten, K. (1999). Steered molecular dynamics simulation of the Rieske subunit motion in the cytochrome bc(1) complex. *Biophys J* **77**, 1753-68.
69. Lovering, F., Aevazelis, C., Chang, J., Dehnhardt, C., Fitz, L., Han, S., Janz, K., Lee, J., Kaila, N., McDonald, J., Moore, W., Moretto, A., Papaioannou, N., Richard, D., Ryan, M. S., Wan, Z. K. & Thorarensen, A. (2016). Imidazotriazines: Spleen Tyrosine Kinase (Syk) Inhibitors Identified by Free-Energy Perturbation (FEP). *ChemMedChem* **11**, 217-33.
70. Barducci, A., Bussi, G. & Parrinello, M. (2008). Well-Tempered Metadynamics: A Smoothly Converging and Tunable Free-Energy Method. *Physical Review Letters* **100**.
71. Hoffmann, T. & Metternich, R. (2012). The future of medicinal chemistry. *Angew Chem Int Ed Engl* **51**, 8670-1.

72. Limongelli, V., Bonomi, M. & Parrinello, M. (2013). Funnel metadynamics as accurate binding free-energy method. *Proc Natl Acad Sci U S A* **110**, 6358-63.
73. Hughes, J. P., Rees, S., Kalindjian, S. B. & Philpott, K. L. (2011). Principles of early drug discovery. *Br J Pharmacol* **162**, 1239-49.
74. Swinney, D. C. & Anthony, J. (2011). How were new medicines discovered? *Nat Rev Drug Discov* **10**, 507-19.
75. Zheng, W., Thorne, N. & McKew, J. C. (2013). Phenotypic screens as a renewed approach for drug discovery. *Drug Discov Today* **18**, 1067-73.
76. Keserű, G. M. & Makara, G. M. (2006). Hit discovery and hit-to-lead approaches. *Drug Discovery Today* **11**, 741-748.
77. Macalino, S. J., Gosu, V., Hong, S. & Choi, S. (2015). Role of computer-aided drug design in modern drug discovery. *Arch Pharm Res* **38**, 1686-701.
78. Lipinski, C. A., Lombardo, F., Dominy, B. W. & Feeney, P. J. (2001). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev* **46**, 3-26.
79. Carr, R. A., Congreve, M., Murray, C. W. & Rees, D. C. (2005). Fragment-based lead discovery: leads by design. *Drug Discov Today* **10**, 987-92.
80. Pellecchia, M., Bertini, I., Cowburn, D., Dalvit, C., Giralt, E., Jahnke, W., James, T. L., Homans, S. W., Kessler, H., Luchinat, C., Meyer, B., Oschkinat, H., Peng, J., Schwalbe, H. & Siegal, G. (2008). Perspectives on NMR in drug discovery: a technique comes of age. *Nat Rev Drug Discov* **7**, 738-45.
81. Congreve, M., Carr, R., Murray, C. & Jhoti, H. (2003). A 'rule of three' for fragment-based lead discovery? *Drug Discov Today* **8**, 876-7.
82. Schuffenhauer, A., Ruedisser, S., Marzinzik, A. L., Jahnke, W., Blommers, M., Selzer, P. & Jacoby, E. (2005). Library design for fragment based screening. *Curr Top Med Chem* **5**, 751-62.
83. Jencks, W. P. (1981). On the attribution and additivity of binding energies. *Proc Natl Acad Sci U S A* **78**, 4046-50.
84. Page, M. I. & Jencks, W. P. (1971). Entropic contributions to rate accelerations in enzymic and intramolecular reactions and the chelate effect. *Proc Natl Acad Sci U S A* **68**, 1678-83.
85. Shuker, S. B., Hajduk, P. J., Meadows, R. P. & Fesik, S. W. (1996). Discovering high-affinity ligands for proteins: SAR by NMR. *Science* **274**, 1531-4.
86. Hajduk, P. J., Dinges, J., Miknis, G. F., Merlock, M., Middleton, T., Kempf, D. J., Egan, D. A., Walter, K. A., Robins, T. S., Shuker, S. B., Holzman, T. F. & Fesik, S. W. (1997). NMR-based discovery of lead inhibitors that block DNA binding of the human papillomavirus E2 protein. *J Med Chem* **40**, 3144-50.
87. Hajduk, P. J., Gerfin, T., Boehlen, J. M., Haberli, M., Marek, D. & Fesik, S. W. (1999). High-throughput nuclear magnetic resonance-based screening. *J Med Chem* **42**, 2315-7.
88. Hajduk, P. J., Meadows, R. P. & Fesik, S. W. (1999). NMR-based screening in drug discovery. *Q Rev Biophys* **32**, 211-40.
89. Heller, M. O., Kessler, H. (2001). NMR Spectroscopy in Drug Design. *Pure Appl. Chem.* **73**, 8.
90. Copeland, R. A., Pompliano, D. L. & Meek, T. D. (2006). Drug-target residence time and its implications for lead optimization. *Nat Rev Drug Discov* **5**, 730-9.

91. Barelrier, S., Linard, D., Pons, J., Clippe, A., Knoops, B., Lancelin, J. M. & Krimm, I. (2010). Discovery of fragment molecules that bind the human peroxiredoxin 5 active site. *PLoS One* **5**, e9744.
92. Hopkins, A. L., Groom, C. R. & Alex, A. (2004). Ligand efficiency: a useful metric for lead selection. *Drug Discov Today* **9**, 430-1.
93. Hajduk, P. J., Meadows, R. P. & Fesik, S. W. (1997). Discovering high-affinity ligands for proteins. *Science* **278**, 497,499.
94. Munos, B. (2009). Lessons from 60 years of pharmaceutical innovation. *Nat Rev Drug Discov* **8**, 959-968.
95. Borhani, D. W. & Shaw, D. E. (2012). The future of molecular dynamics simulations in drug discovery. *J Comput Aided Mol Des* **26**, 15-26.
96. Trott, O. & Olson, A. J. (2010). AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem* **31**, 455-61.
97. Aparoy, P., Reddy, K. K. & Reddanna, P. (2012). Structure and ligand based drug design strategies in the development of novel 5- LOX inhibitors. *Curr Med Chem* **19**, 3763-78.
98. ARCHER Website.
99. Haohuan, F., Junfeng, L., Jinzhe, Y., Lanning, W., Xiaomeng, H., Chao, Y., Wei, X., Fangli, Q., Wei, Z., Xunqiang, Y., Chaofeng, H., Wei, G., Jian, Z., Yangang, W. & Guangwen, Y. (2016). The Sunway TaihuLight supercomputer: \\system and applications. *SCIENCE CHINA Information Sciences* **59**, 072001.
100. Zhao, H. & Caflisch, A. (2015). Molecular dynamics in drug design. *Eur J Med Chem* **91**, 4-14.
101. Ignasi Buch, T. G., and Gianni De Fabritiis. (2011). Complete reconstruction of an enzyme-inhibitor binding process by molecular dynamics simulations. *Proc Natl Acad Sci U S A* **108**, 6.
102. Poynton, R. A. & Hampton, M. B. (2014). Peroxiredoxins as biomarkers of oxidative stress. *Biochim Biophys Acta* **1840**, 906-12.
103. Perkins, A., Nelson, K. J., Parsonage, D., Poole, L. B. & Karplus, P. A. (2015). Peroxiredoxins: guardians against oxidative stress and modulators of peroxide signaling. *Trends Biochem Sci* **40**, 435-45.
104. Winterbourn, C. C. (2008). Reconciling the chemistry and biology of reactive oxygen species. *Nat Chem Biol* **4**, 278-86.
105. Perkins, A., Poole, L. B. & Karplus, P. A. (2014). Tuning of peroxiredoxin catalysis for various physiological roles. *Biochemistry* **53**, 7693-705.
106. Hall, A., Parsonage, D., Poole, L. B. & Karplus, P. A. (2010). Structural evidence that peroxiredoxin catalytic power is based on transition-state stabilization. *J Mol Biol* **402**, 194-209.
107. Barranco-Medina, S., Lazaro, J. J. & Dietz, K. J. (2009). The oligomeric conformation of peroxiredoxins links redox state to function. *FEBS Lett* **583**, 1809-16.
108. Rhee, S. G., Kang, S. W., Chang, T. S., Jeong, W. & Kim, K. (2001). Peroxiredoxin, a novel family of peroxidases. *IUBMB Life* **52**, 35-41.
109. Wood, Z. A., Poole, L. B. & Karplus, P. A. (2003). Peroxiredoxin evolution and the regulation of hydrogen peroxide signaling. *Science* **300**, 650-3.
110. Knoops, B., Argyropoulou, V., Becker, S., Ferte, L. & Kuznetsova, O. (2016). Multiple Roles of Peroxiredoxins in Inflammation. *Mol Cells* **39**, 60-4.

111. Nelson, K. J., Knutson, S. T., Soito, L., Klomsiri, C., Poole, L. B. & Fetrow, J. S. (2011). Analysis of the peroxiredoxin family: using active-site structure and sequence information for global classification and residue analysis. *Proteins* **79**, 947-64.
112. Flohe, L. & Harris, J. R. (2007). Introduction. History of the peroxiredoxins and topical perspectives. *Subcell Biochem* **44**, 1-25.
113. Parsonage, D., Nelson, K. J., Ferrer-Sueta, G., Alley, S., Karplus, P. A., Furdui, C. M. & Poole, L. B. (2015). Dissecting peroxiredoxin catalysis: separating binding, peroxidation, and resolution for a bacterial AhpC. *Biochemistry* **54**, 1567-75.
114. Rao, S. T. & Rossmann, M. G. (1973). Comparison of super-secondary structures in proteins. *Journal of Molecular Biology* **76**, 241-256.
115. Declercq, J. P., Evrard, C., Clippe, A., Stricht, D. V., Bernard, A. & Knoops, B. (2001). Crystal structure of human peroxiredoxin 5, a novel type of mammalian peroxiredoxin at 1.5 Å resolution. *J Mol Biol* **311**, 751-9.
116. Mizohata, E., Sakai, H., Fusatomi, E., Terada, T., Murayama, K., Shirouzu, M. & Yokoyama, S. (2005). Crystal structure of an archaeal peroxiredoxin from the aerobic hyperthermophilic crenarchaeon *Aeropyrum pernix* K1. *J Mol Biol* **354**, 317-29.
117. Knoops, B., Goemaere, J., Van der Eecken, V. & Declercq, J. P. (2011). Peroxiredoxin 5: structure, mechanism, and function of the mammalian atypical 2-Cys peroxiredoxin. *Antioxid Redox Signal* **15**, 817-29.
118. Aran, M., Ferrero, D., Wolosiuk, A., Mora-García, S. & Wolosiuk, R. A. (2011). ATP and Mg(2+) Promote the Reversible Oligomerization and Aggregation of Chloroplast 2-Cys Peroxiredoxin. *The Journal of Biological Chemistry* **286**, 23441-23451.
119. Claiborne, A., Yeh, J. I., Mallett, T. C., Luba, J., Crane, E. J., 3rd, Charrier, V. & Parsonage, D. (1999). Protein-sulfenic acids: diverse roles for an unlikely player in enzyme catalysis and redox regulation. *Biochemistry* **38**, 15407-16.
120. Rhee, S. G., Chae, H. Z. & Kim, K. (2005). Peroxiredoxins: a historical overview and speculative preview of novel mechanisms and emerging concepts in cell signaling. *Free Radic Biol Med* **38**, 1543-52.
121. Fujii, J. & Ikeda, Y. (2002). Advances in our understanding of peroxiredoxin, a multifunctional, mammalian redox protein. *Redox Rep* **7**, 123-30.
122. Knoops, B., Clippe, A., Bogard, C., Arsalane, K., Wattiez, R., Hermans, C., Duconseille, E., Falmagne, P. & Bernard, A. (1999). Cloning and characterization of AOEB166, a novel mammalian antioxidant enzyme of the peroxiredoxin family. *J Biol Chem* **274**, 30451-8.
123. Bryk, R., Griffin, P. & Nathan, C. (2000). Peroxynitrite reductase activity of bacterial peroxiredoxins. *Nature* **407**, 211-5.
124. Dalvit, C., Fogliatto, G., Stewart, A., Veronesi, M. & Stockman, B. (2001). WaterLOGSY as a method for primary NMR screening: practical aspects and range of applicability. *J Biomol NMR* **21**, 349-59.
125. Parsonage, D., Karplus, P. A. & Poole, L. B. (2008). Substrate specificity and redox potential of AhpC, a bacterial peroxiredoxin. *Proc Natl Acad Sci U S A* **105**, 8209-14.
126. Kang, S. W., Rhee, S. G., Chang, T. S., Jeong, W. & Choi, M. H. (2005). 2-Cys peroxiredoxin function in intracellular signal transduction: therapeutic implications. *Trends Mol Med* **11**, 571-8.
127. Reed, T. T., Pierce, W. M., Jr., Turner, D. M., Markesbery, W. R. & Butterfield, D. A. (2009). Proteomic identification of nitrated brain proteins in early Alzheimer's disease inferior parietal lobule. *J Cell Mol Med* **13**, 2019-29.

128. Holmgren, A. (1995). Thioredoxin structure and mechanism: conformational changes on oxidation of the active-site sulfhydryls to a disulfide. *Structure* **3**, 239-243.
129. Rhee, S. G., Jeong, W., Chang, T. S. & Woo, H. A. (2007). Sulfiredoxin, the cysteine sulfinic acid reductase specific to 2-Cys peroxiredoxin: its discovery, mechanism of action, and biological significance. *Kidney Int Suppl*, S3-8.
130. Karplus, P. A. (2015). A primer on peroxiredoxin biochemistry. *Free Radic Biol Med* **80**, 183-90.
131. Smeets, A., Marchand, C., Linard, D., Knoop, B. & Declercq, J. P. (2008). The crystal structures of oxidized forms of human peroxiredoxin 5 with an intramolecular disulfide bond confirm the proposed enzymatic mechanism for atypical 2-Cys peroxiredoxins. *Arch Biochem Biophys* **477**, 98-104.
132. Evrard, C., Capron, A., Marchand, C., Clippe, A., Wattiez, R., Soumillion, P., Knoop, B. & Declercq, J. P. (2004). Crystal structure of a dimeric oxidized form of human peroxiredoxin 5. *J Mol Biol* **337**, 1079-90.
133. Nicolussi, A., D'Inzeo, S., Capalbo, C., Giannini, G. & Coppa, A. (2017). The role of peroxiredoxins in cancer. *Molecular and Clinical Oncology* **6**, 139-153.
134. Lakhan, S. E., Kirchgessner, A. & Hofer, M. (2009). Inflammatory mechanisms in ischemic stroke: therapeutic approaches. *J Transl Med* **7**, 97.
135. Moskowitz, M. A., Lo, E. H. & Iadecola, C. (2010). The science of stroke: mechanisms in search of treatments. *Neuron* **67**, 181-98.
136. Macrez, R., Ali, C., Toutirais, O., Le Mauff, B., Defer, G., Dirnagl, U. & Vivien, D. (2011). Stroke and the immune system: from pathophysiology to new therapeutic strategies. *Lancet Neurol* **10**, 471-80.
137. INSERM. (2013). AVC - Accident vasculaire Cerebral.
138. Garcia-Bonilla, L. & Iadecola, C. (2012). Peroxiredoxin sets the brain on fire after stroke. *Nat Med* **18**, 858-9.
139. Lawes, C. M., Bennett, D. A., Feigin, V. L. & Rodgers, A. (2004). Blood pressure and stroke: an overview of published reviews. *Stroke* **35**, 1024.
140. O'Donnell, M. J., Xavier, D., Liu, L., Zhang, H., Chin, S. L., Rao-Melacini, P., Rangarajan, S., Islam, S., Pais, P., McQueen, M. J., Mondo, C., Damasceno, A., Lopez-Jaramillo, P., Hankey, G. J., Dans, A. L., Yusuf, S., Truelsen, T., Diener, H. C., Sacco, R. L., Ryglewicz, D., Czlonkowska, A., Weimar, C., Wang, X., Yusuf, S. & investigators, I. (2010). Risk factors for ischaemic and intracerebral haemorrhagic stroke in 22 countries (the INTERSTROKE study): a case-control study. *Lancet* **376**, 112-23.
141. Balami, J. S., Chen, R. L., Grunwald, I. Q. & Buchan, A. M. (2011). Neurological complications of acute ischaemic stroke. *Lancet Neurol* **10**, 357-71.
142. Kumar S., S. M. H., Caplan L. R. (2010). Medical complications after stroke. *Lancet Neurol* **9**, 14.
143. del Zoppo, G. J. (2010). Plasminogen activators in ischemic stroke: introduction. *Stroke* **41**, S39-41.
144. Shichita, T., Sakaguchi, R., Suzuki, M. & Yoshimura, A. (2012). Post-ischemic inflammation in the brain. *Front Immunol* **3**, 132.
145. Le Behot, A., Gauberti, M., Martinez De Lizarrondo, S., Montagne, A., Lemarchand, E., Repesse, Y., Guillou, S., Denis, C. V., Maubert, E., Orset, C. & Vivien, D. (2014). GpIbalpha-VWF blockade restores vessel patency by dissolving platelet aggregates formed under very high shear rate in mice. *Blood* **123**, 3354-63.

146. Ejaz, S., Emmrich, J. V., Sitnikov, S. L., Hong, Y. T., Sawiak, S. J., Fryer, T. D., Aigbirhio, F. I., Williamson, D. J. & Baron, J. C. (2016). Normobaric hyperoxia markedly reduces brain damage and sensorimotor deficits following brief focal ischaemia. *Brain* **139**, 751-64.
147. Shichita, T., Hasegawa, E., Kimura, A., Morita, R., Sakaguchi, R., Takada, I., Sekiya, T., Ooboshi, H., Kitazono, T., Yanagawa, T., Ishii, T., Takahashi, H., Mori, S., Nishibori, M., Kuroda, K., Akira, S., Miyake, K. & Yoshimura, A. (2012). Peroxiredoxin family proteins are key initiators of post-ischemic inflammation in the brain. *Nat Med* **18**, 911-7.
148. Samson, Y., Lapergue, B. & Hosseini, H. (2005). [Inflammation and ischaemic stroke: current status and future perspectives]. *Rev Neurol (Paris)* **161**, 1177-82.
149. Iadecola, C. & Anrather, J. (2011). The immunology of stroke: from mechanisms to translation. *Nat Med* **17**, 796-808.
150. Copeland, R. A. (2005). *Evaluation of enzyme inhibitors in drug discovery : a guide for medicinal chemists and pharmacologists.*, John Wiley & Sons, New Jersey.
151. Cornish-Bowden, A. J., M.; Saks. V. (2005). *Cinétique Enzymatique*, ESP Sciences, Paris.
152. Cornish-Bowden, A. W., C. W. (1988). *Enzyme Kinetics*, IRL Press, Oxford.
153. Segel, I. H. (1993). *Enzyme Kinetics : Behavior and Analysis of Rapid Equilibrium and Steady-State Enzyme Systems*, Wiley-Intersciences.
154. Portillo-Ledesma, S., Sardi, F., Manta, B., Tourn, M. V., Clippe, A., Knoop, B., Alvarez, B., Coitino, E. L. & Ferrer-Sueta, G. (2014). Deconstructing the catalytic efficiency of peroxiredoxin-5 peroxidatic cysteine. *Biochemistry* **53**, 6113-25.
155. Wolfenden, R. & Snider, M. J. (2001). The depth of chemical time and the power of enzymes as catalysts. *Acc Chem Res* **34**, 938-45.
156. Yang, Y. J., Baek, J. Y., Goo, J., Shin, Y., Park, J. K., Jang, J. Y., Wang, S. B., Jeong, W., Lee, H. J., Um, H. D., Lee, S. K., Choi, Y., Rhee, S. G. & Chang, T. S. (2016). Effective Killing of Cancer Cells Through ROS-Mediated Mechanisms by AMRI-59 Targeting Peroxiredoxin I. *Antioxid Redox Signal* **24**, 453-69.
157. Wei, W., Ma, C., Cao, Y., Yang, L., Huang, Z., Qin, D., Chen, Y., Liu, C., Xia, L., Wang, T., Lei, H., Yu, Y., Huang, M., Tong, Y., Xu, H., Gao, F., Zhang, J. & Wu, Y. L. (2016). Identification of H7 as a novel peroxiredoxin I inhibitor to induce differentiation of leukemia cells. *Oncotarget* **7**, 3873-83.
158. Hansen, J., Palmfeldt, J., Vang, S., Corydon, T. J., Gregersen, N. & Bross, P. (2011). Quantitative proteomics reveals cellular targets of celastrol. *PLoS One* **6**, e26634.
159. Zhao, Q., Ding, Y., Deng, Z., Lee, O.-Y., Gao, P., Chen, P., Rose, R. J., Zhao, H., Zhang, Z., Tao, X.-P., Heck, A. J. R., Kao, R. & Yang, D. (2015). Natural products triptolide, celastrol, and withaferin A inhibit the chaperone activity of peroxiredoxin I. *Chemical Science* **6**, 4124-4130.
160. Liu, G., Botting, C. H., Evans, K. M., Walton, J. A., Xu, G., Slawin, A. M. & Westwood, N. J. (2010). Optimisation of conoidin A, a peroxiredoxin inhibitor. *ChemMedChem* **5**, 41-5.
161. Haraldsen, J. D., Liu, G., Botting, C. H., Walton, J. G., Storm, J., Phalen, T. J., Kwok, L. Y., Soldati-Favre, D., Heintz, N. H., Muller, S., Westwood, N. J. & Ward, G. E. (2009). Identification of Conoidin A as a Covalent Inhibitor of Peroxiredoxin II. *Org Biomol Chem* **7**, 3040-3048.
162. Liu, C. X., Yin, Q. Q., Zhou, H. C., Wu, Y. L., Pu, J. X., Xia, L., Liu, W., Huang, X., Jiang, T., Wu, M. X., He, L. C., Zhao, Y. X., Wang, X. L., Xiao, W. L., Chen, H. Z., Zhao, Q., Zhou, A. W., Wang, L. S., Sun, H. D. & Chen, G. Q. (2012). Adenanthin targets peroxiredoxin I and II to induce differentiation of leukemic cells. *Nat Chem Biol* **8**, 486-93.

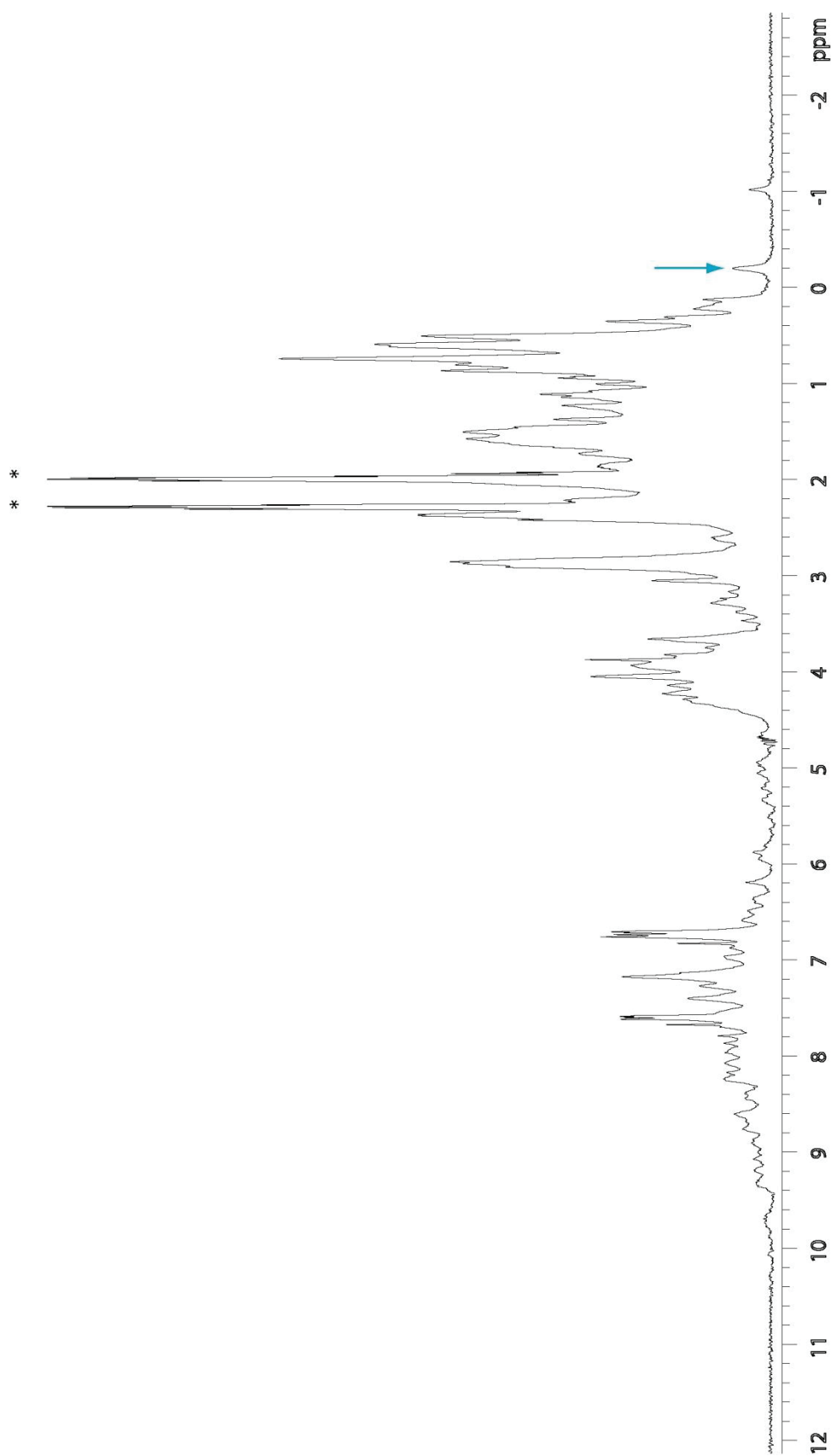
163. Schweigert, N., Zehnder, A. J. & Eggen, R. I. (2001). Chemical properties of catechols and their molecular modes of toxic action in cells, from microorganisms to mammals. *Environ Microbiol* **3**, 81-91.
164. Zheng, L. T., Ryu, G. M., Kwon, B. M., Lee, W. H. & Suk, K. (2008). Anti-inflammatory effects of catechols in lipopolysaccharide-stimulated microglia cells: inhibition of microglial neurotoxicity. *Eur J Pharmacol* **588**, 106-13.
165. Humphrey, W., Dalke, A. & Schulten, K. (1996). VMD: visual molecular dynamics. *J Mol Graph* **14**, 33-8, 27-8.
166. Williams, T. & Kelley, C. (2011). Gnuplot 4.4 : An Interactive Plotting Program.
167. Hanwell, M. D., Curtis, D. E., Lonie, D. C., Vandermeersch, T., Zurek, E. & Hutchison, G. R. (2012). Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *J Cheminform* **4**, 17.
168. Wang, J., Wang, W., Kollman, P. A. & Case, D. A. (2006). Automatic atom type and bond type perception in molecular mechanical calculations. *J Mol Graph Model* **25**, 247-60.
169. Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A. & Case, D. A. (2004). Development and testing of a general amber force field. *J Comput Chem* **25**, 1157-74.
170. Schroder, E., Littlechild, J. A., Lebedev, A. A., Errington, N., Vagin, A. A. & Isupov, M. N. (2000). Crystal structure of decameric 2-Cys peroxiredoxin from human erythrocytes at 1.7 Å resolution. *Structure* **8**, 605-15.
171. Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S. & Olson, A. J. (2009). AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. *Journal of computational chemistry* **30**, 2785-2791.
172. Harwood, C. R. e. a. (1990). Chemically defined growth media and supplements. In *Molecular Biological methods for Bacillus* (Harwood, C. R. a. C., S. M., ed.), pp. 548. Wiley, Chichester.
173. Tran, P. H., Duus, F. & Le, T. N. (2012). Friedel–Crafts acylation using bismuth triflate in [BMI][PF₆]. *Tetrahedron Letters* **53**, 222-224.
174. Guyon, C., Baron, M., Lemaire, M., Popowycz, F. & Méta y, E. (2014). Commutative reduction of aromatic ketones to arylmethylenes/alcohols by hypophosphites catalyzed by Pd/C under biphasic conditions. *Tetrahedron* **70**, 2088-2095.
175. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J. & Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* **6**, 277-93.
176. Johnson, B. A. & Blevins, R. A. (1994). NMR View: A computer program for the visualization and analysis of NMR data. *J Biomol NMR* **4**, 603-14.
177. Fielding, L. (2007). NMR methods for the determination of protein–ligand dissociation constants. *Progress in Nuclear Magnetic Resonance Spectroscopy* **51**, 219-242.
178. Lange, A., Ismail, M. B., Riviere, G., Hologne, M., Lacabanne, D., Guilliere, F., Lancelin, J. M., Krimm, I. & Walker, O. (2012). Competitive binding of UBPY and ubiquitin to the STAM2 SH3 domain revealed by NMR. *FEBS Lett* **586**, 3379-84.
179. Kim, J. A., Park, S., Kim, K., Rhee, S. G. & Kang, S. W. (2005). Activity assay of mammalian 2-cys peroxiredoxins using yeast thioredoxin reductase system. *Anal Biochem* **338**, 216-23.
180. Kuzmic, P. (1996). Program DYNAFIT for the analysis of enzyme kinetic data: application to HIV proteinase. *Anal Biochem* **237**, 260-73.

181. Kuzmic, P. (2009). DynaFit--a software package for enzymology. *Methods Enzymol* **467**, 247-80.
182. Aguirre, C., ten Brink, T., Guichou, J. F., Cala, O. & Krimm, I. (2014). Comparing binding modes of analogous fragments using NMR in fragment-based drug design: application to PRDX5. *PLoS One* **9**, e102300.
183. Troussicot, L., Guilliere, F., Limongelli, V., Walker, O. & Lancelin, J. M. (2015). Funnel-metadynamics and solution NMR to estimate protein-ligand affinities. *J Am Chem Soc* **137**, 1273-81.
184. Otting, G., Liepinsh, E., Farmer, B. T., 2nd & Wuthrich, K. (1991). Protein hydration studied with homonuclear 3D 1H NMR experiments. *J Biomol NMR* **1**, 209-15.
185. Cardey, B. & Enescu, M. (2005). A Computational Study of Thiolate and Selenolate Oxidation by Hydrogen Peroxide. *ChemPhysChem* **6**, 1175-1180.
186. Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J. L., Dror, R. O. & Shaw, D. E. (2010). Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins: Structure, Function, and Bioinformatics*, NA-NA.
187. Buch, I., Giorgino, T. & De Fabritiis, G. (2011). Complete reconstruction of an enzyme-inhibitor binding process by molecular dynamics simulations. *Proc Natl Acad Sci U S A* **108**, 10184-9.
188. Chow, M. L., Troussicot, L., Martin, M., Doumeche, B., Guilliere, F. & Lancelin, J. M. (2016). Predicting and Understanding the Enzymatic Inhibition of Human Peroxiredoxin 5 by 4-Substituted Pyrocatechols by Combining Funnel Metadynamics, Solution NMR, and Steady-State Kinetics. *Biochemistry* **55**, 3469-80.
189. Sinnokrot, M. O. & Sherrill, C. D. (2006). High-Accuracy Quantum Mechanical Studies of π - π Interactions in Benzene Dimers. *The Journal of Physical Chemistry A* **110**, 10656-10668.
190. Trujillo, M., Clippe, A., Manta, B., Ferrer-Sueta, G., Smeets, A., Declercq, J. P., Knoops, B. & Radi, R. (2007). Pre-steady state kinetic characterization of human peroxiredoxin 5: taking advantage of Trp84 fluorescence increase upon oxidation. *Arch Biochem Biophys* **467**, 95-106.
191. Kim, K. H., Lee, W. & Kim, E. E. (2016). Crystal structures of human peroxiredoxin 6 in different oxidation states. *Biochemical and Biophysical Research Communications* **477**, 717-722.

ANNEXES

ANNEXE 1. SPECTRE 1D DE HPRDX5 A PH=7.4	219
ANNEXE 2. DEMONSTRATION DE L'ÉQUATION 35 RELIANT LA PERTURBATION DE DEPLACEMENT CHIMIQUE CSP A LA CONSTANTE DE DISSOCIATION K_D D'UN COMPLEXE PROTEINE LIGAND PAR RMN	220
ANNEXE 3. TABLEAU RECAPITULATIF DES FORMULES ET PROPRIETES CHIMIQUES DES LIGANDS ETUDIES.....	221
ANNEXE 4. SYNTHÈSE DU 4NPC : CARACTERISATION DES INTERMÉDIAIRES DE SYNTHÈSE ET DU PRODUIT FINAL 4NPC.	223
ANNEXE 5. SPECTRES RMN DE CARACTERISATION DU PRODUIT 4NPC SYNTHÉTISÉ	224
ANNEXE 6. PARAMÈTRES UTILISÉS PAR LE LOGICIEL DYNAFIT POUR LA DÉTERMINATION DU MÉCANISME D'INHIBITION ET DES CONSTANTES ASSOCIÉES.	226
ANNEXE 7. COMPARAISON DES VALEURS D'IC ₅₀ OBTENUS POUR [H ₂ O ₂] = K _M ET [H ₂ O ₂] = 10K _M	229

Annexe 1. Spectre 1D 1H de hPRDX5 200 μ M à 600 MHz (sonde cryo) dans un tampon PBS à pH=7.4, et 300K. Protéine maintenue sous forme réduite par 3mM de TCEP (signaux du TCEP signifiés par *). Signal irradié sélectivement lors des expériences STD indiqué par une flèche.



Annexe 2. Démonstration de l'équation 35 reliant la perturbation de déplacement chimique CSP à la constante de dissociation K_D d'un complexe protéine ligand par RMN

Si l'on considère une réaction entre une protéine P et un ligand L pour donner de manière réversible un complexe protéine-ligand PL, à l'équilibre on peut définir la constante de dissociation du complexe protéine-ligand K_D .



A l'équilibre:
$$K_D = \frac{[P]_t[L]_t}{[PL]_t}$$

Avec $[P]_t = [P]_0 - [PL]_t$ et $[L]_t = [L]_0 - [PL]_t$

En remplaçant dans l'expression de K_D on obtient alors la relation suivante :

$$K_D = \frac{([P]_0 - [PL]_t)([L]_0 - [PL]_t)}{[PL]_t} = \frac{[P]_0[L]_0 - ([P]_0 + [L]_0)[PL]_t + [PL]_t^2}{[PL]_t}$$

Soit $K_D [PL]_t = [P]_0[L]_0 - ([P]_0 + [L]_0)[PL]_t + [PL]_t^2$

Ce qui nous donne l'équation du second degré suivante :

$$[PL]_t^2 - ([P]_0 + [L]_0)[PL]_t - K_D [PL]_t + [P]_0[L]_0 = 0$$

En considérant l'équation du second degré : $ax^2 + bx + c = 0$ et en appliquant la formule quadratique,

on a : $x = \frac{-b \pm \sqrt{\Delta}}{2a}$ avec $\Delta = b^2 - 4ac$ le discriminant

Soit:
$$[PL]_t = \frac{([P]_0 + [L]_0 + [K_D]) \pm \sqrt{([P]_0 + [L]_0 + [K_D])^2 - 4([P]_0[L]_0)}}{2a}$$

En considérant un modèle 1 :1 où un seul ligand vient se lier dans un seul site actif du récepteur, on peut écrire que :

$$\Delta\delta_{obs} = \frac{[PL]_t \cdot \Delta\delta_{max}}{[P]_0}$$

Avec $\Delta\delta_{max}$ le déplacement chimique maximum, observé à saturation en ligand.

En combinant à présent la solution de l'équation du second degré avec la relation propre à notre modèle 1:1, on obtient la relation suivante reliant la perturbation de déplacement chimique observé, la constante de dissociation du complexe PL et les concentrations initiales en protéine $[P]_0$ et en ligand $[L]_0$:

$$\Delta\delta_{obs} = \Delta\delta_{max} \cdot \frac{([P]_0 + [L]_0 + K_D) - \sqrt{([P]_0 + [L]_0 + K_D)^2 - (4[P]_0[L]_0)}}{2[P]_0}$$

Annexe 3. Tableau récapitulatif des formules et propriétés chimiques des ligands étudiés

Molécule	Formule Brute	Formule SMILES	MW (g/mol)	logP ^s	LE*	Atomes Lourds	Donneurs LH (NH/OH)	Accepteurs LH (O/N)
catéchol	C6H6O2	Oc1ccccc1O	110.11	0.99	0.38	8	2	2
cyclohexanediol- <i>cis</i>	C6H12O2	C1CCC(C(C1)O)O	116.16	0.15	0.29	8	2	2
cyclohexanediol- <i>trans</i>	C6H12O2	C1CCC(C(C1)O)O	116.16	0.15	0.30	8	2	2
4-méthylcatéchol	C7H8O2	Cc1ccc(O)c(O)c1	124.14	1.42	0.47	9	2	2
4- <i>tert</i> -butylcatéchol	C10H14O2	CC(C)(C)c1ccc(O)c(O)c1	166.22	2.67	0.46	12	2	2
4- <i>iso</i> -propylcatéchol	C9H12O2	CC(C)c1ccc(O)c(O)c1	152.19	2.48	0.60	11	2	2
4- <i>iso</i> -butylcatéchol	C10H14O2	CCC(C)c1ccc(O)c(O)c1	166.22	2.83	0.36	12	2	2
4-éthylcatéchol	C8H10O2	Ccc1ccc(O)c(O)c1	138.17	1.88	0.36	10	2	2
gallate d'éthyle	C9H10O5	CCOC(=O)c1cc(O)c(O)c(O)c1	198.17	1.23	0.40	14	3	5
acide (3,4-dihydroxyphényl) acétique	C8H8O4	OC(=O)Cc1ccc(O)c(O)c1	168.15	0.39	0.48	12	3	4
4-aminométhylcatéchol	C7H9NO2	Oc1ccc(CN)cc1O	139.15	0.15	0.21	10	4	3
hydroperoxyde de cumène	C9H12O2	CC(C)(OO)c1ccccc1	152.19	2.38	0.19	11	1	2
2-phénylpropan-2-ol	C9H12O	CC(O)(C)c1ccccc1	136.19	2.08	0.19	10	1	1
hydroperoxyde de benzène	C7H8O2	OOc1ccccc1	124.14	1.57	0.28	9	1	2

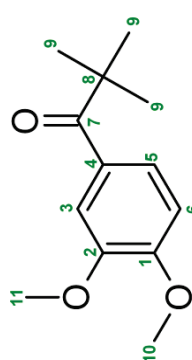
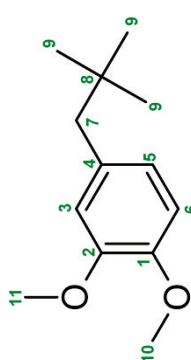
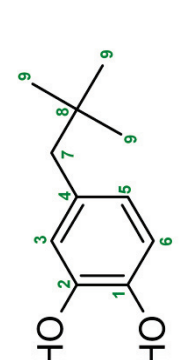
4-néo-pentylcatéchol	C11H16O2	CC(C)(C)Cc1ccc(O)c(O)c1	180.247	2.85	0.47	13	2	2
6-(néo-pentyl)biphényl-3,4-diol	C17H20O2	Oc2cc(c1ccccc1)c(cc2O)CC(C)(C)C	256.34	4.57	0.36	19	2	2
4-benzyl-5-(néo-pentyl)catéchol	C18H22O2	Oc2cc(Cc1ccccc1)c(cc2O)CC(C)(C)C	270.37	4.74	0.32	20	2	2
4-(néo-pentyl)-5-(phénylamino)catéchol	C17H21NO2	Oc1cc(CC(C)(C)C)c(cc1O)Nc2ccccc2	271.35	4.77	0.41	20	3	3
4-[(4-aminophényl)amino]-5-(néo-pentyl)catéchol	C17H22N2O2	Oc1cc(CC(C)(C)C)c(cc1O)Nc2ccc(N)c2	286.37	3.85	0.27	21	5	4
4-[[2-(néo-pentyl)-4,5-dihydroxyphényl]amino]benzoate	C18H20NO4	Oc1cc(CC(C)(C)C)c(cc1O)Nc2ccc(cc2)C([O-])=O	314.36	1.27	0.21	23	3	5
4-[[2-(néo-pentyl)-4,5-dihydroxyphényl]amino]phényl)méthaniminum	C18H25N2O2	Oc1cc(CC(C)(C)C)c(cc1O)Nc2ccc(C[NH3+])cc2	301.40	1.41	0.37	22	6	4

* LE : Efficacité ligand, calculée comme étant le rapport de l'énergie libre de liaison du complexe (ΔG^\ddagger) sur le nombre d'atomes lourds du ligand

[§]log P : coefficient de partition octanol/eau prédit par le logiciel Molinspiration Property Calculation Service de la compagnie Molinspiration Cheminformatics

Annexe 4. Synthèse du 4NPC : Caractérisation des intermédiaires de synthèse et du produit final 4NPC.

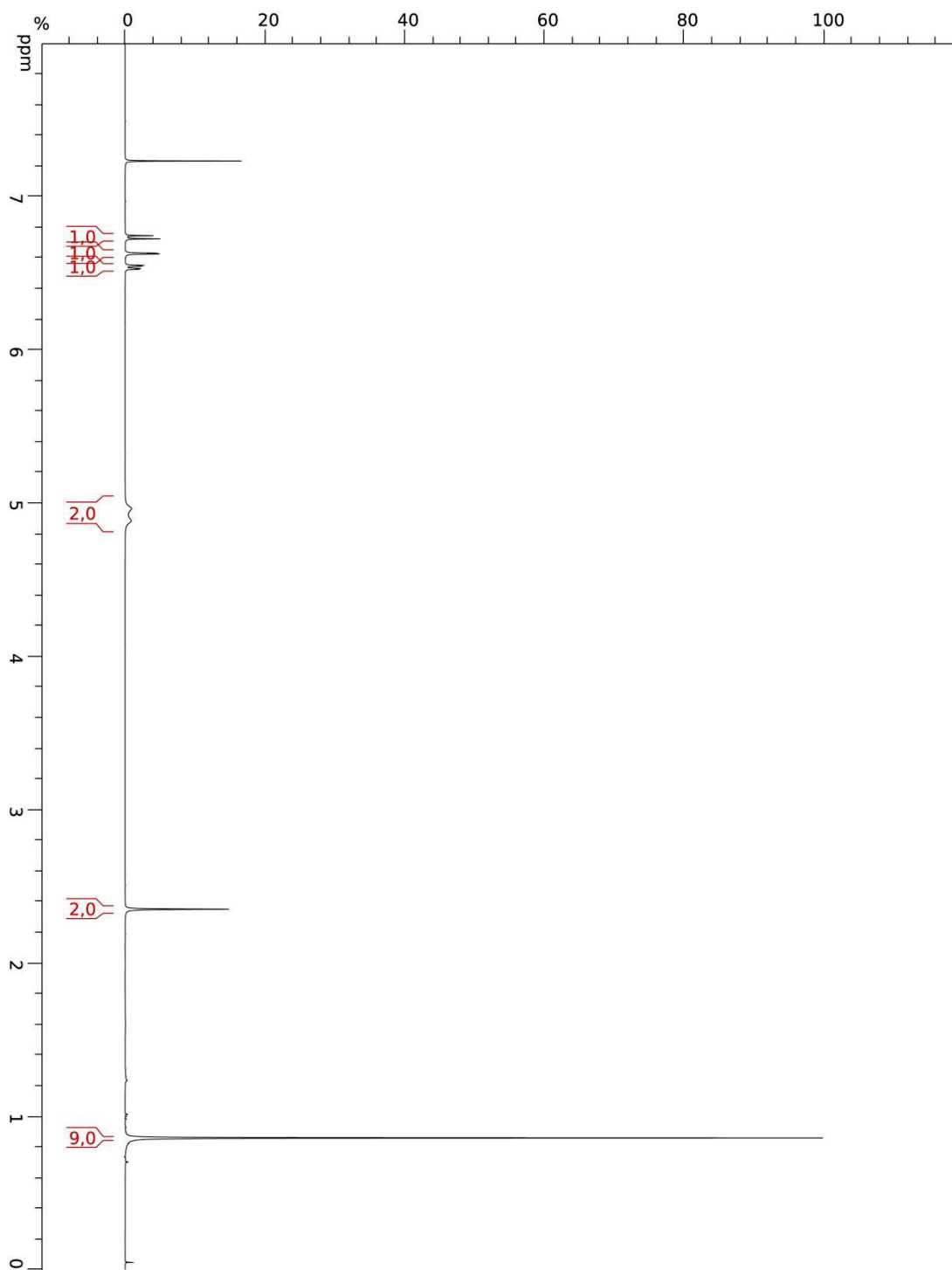
Par : CCM (R_f), RMN (^1H et ^{13}C), ESI-MS (valeur théorique calculée et valeur trouvée expérimentalement) et point de fusion.

Molécule	R_f	Tf (°C)	RMN ^1H (ppm)	RMN ^{13}C (ppm)	ESI-MS
	DCM : EP* 3 : 2 0.21	-	1.33 (9H, s, CMe ₃) 3.86 (3H, s, OMe) 3.87 (3H, s, OMe) 6.79 (1H, d, J_{6-5} 8.5 Hz, H ₆) 7.36 (1H, d, J_{3-5} 2 Hz, H ₃) 7.48 (1H, dd, J_{5-6} 8.5 Hz, J_{5-3} 2 Hz, H ₅)	28.7 (CMe ₃), 44.1 (CMe ₃), 56.06 (OMe), 56.08 (OMe), 109.8 (CH _{ar}), 112.3 (CH _{ar}), 122.8 (CH _{ar}), 130.2 (Cq _{ar}), 148.8 (Cq _{ar} O), 160.0 (Cq _{ar} O), 206.2 (C=O)	[M+Na] ⁺ Calc : 245.1148 Exp : 245.1145
	DCM : EP 7 : 3 0.77	36-37	0.89 (9H, s, CMe ₃) 2.42 (2H, s, CH ₂) 3.84 (6H, s, OMe) 6.63 (1H, s, H ₃) 6.65 (1H, dd, J_{5-6} 8 Hz, J_{5-3} 2 Hz, H ₅) 6.76 (1H, d, J_{6-5} 8 Hz, H ₆)	29.6 (CMe ₃), 32.0 (CMe ₃), 50.1 (CH ₂), 56.08 (OMe), 56.09 (OMe), 110.9 (CH _{ar}), 114.3 (CH _{ar}), 122.7 (CH _{ar}), 132.7 (Cq _{ar}), 147.5 (Cq _{ar} O), 148.4 (Cq _{ar} O)	[M+Na] ⁺ Calc : 231.1356 Exp : 231.1351
	DCM 0.17	124-125	0.86 (9H, s, CMe ₃) 2.35 (2H, s, CH ₂) 4.88 (1H, s, OH) 4.96 (1H, s, OH) 6.54 (1H, dd, J_{5-6} 8 Hz, J_{5-3} 2 Hz, H ₅), 6.62 (1H, d, J_{3-5} 2 Hz, H ₃) 6.73 (1H, d, J_{6-5} 8 Hz, H ₆)	29.6 (CMe ₃), 31.9 (CMe ₃), 50.0 (CH ₂), 114.9 (CH _{ar}), 117.8 (CH _{ar}), 123.3 (CH _{ar}), 133.4 (Cq _{ar}), 141.7 (Cq _{ar} O), 142.9 (Cq _{ar} O)	[M-H] ⁻ Calc : 179.1078 Exp : 179.1071

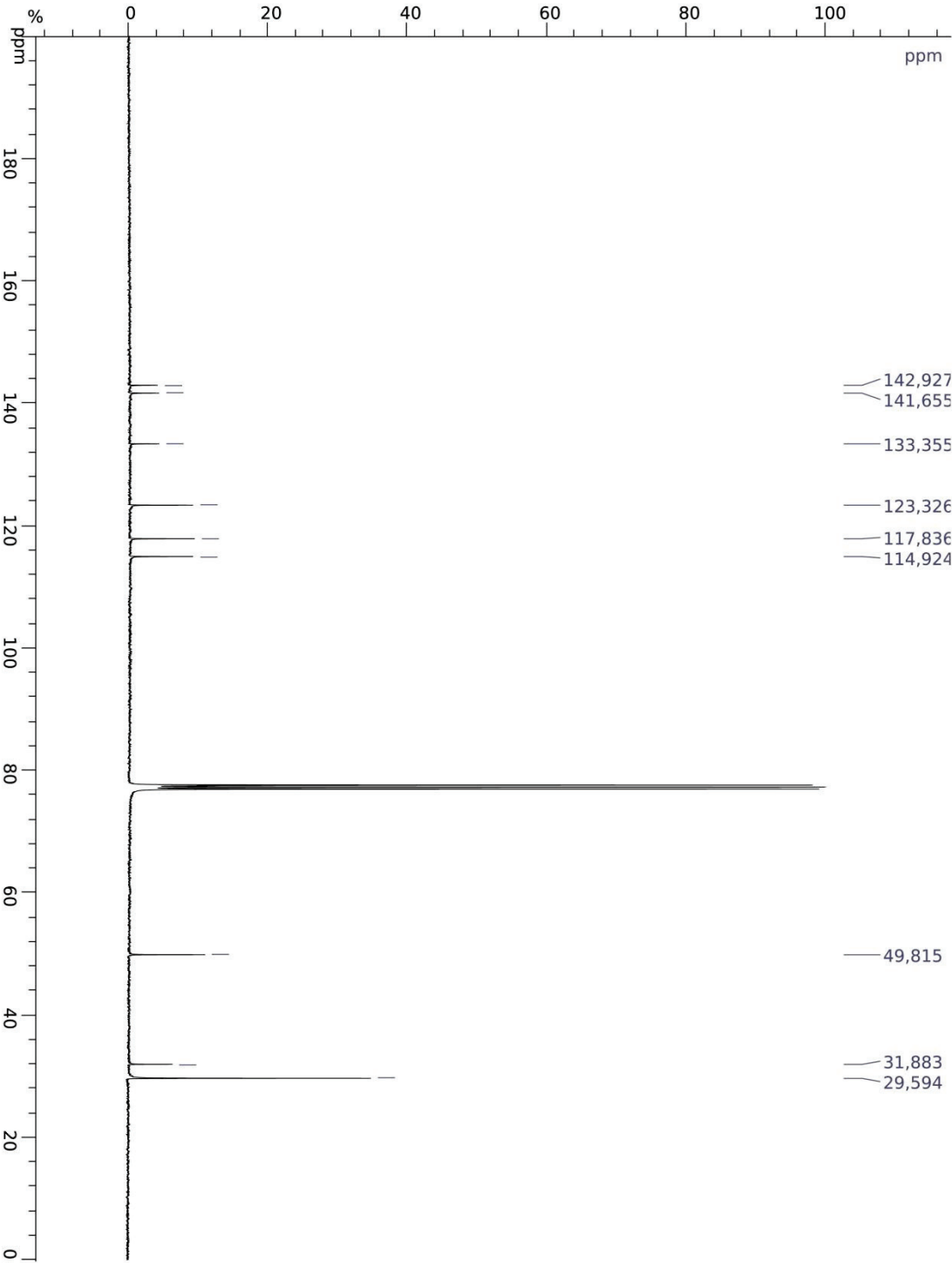
*DCM : Dichlorométhane ; EP : Ether de pétrole

Annexe 5. Spectres RMN de caractérisation du produit 4NPC synthétisé

A) Spectre ^1H



B) Spectre ¹³C



Annexe 6. Paramètres utilisés par le logiciel DynaFit pour la détermination du mécanisme d'inhibition et des constantes associées.

Les algorithmes de Levenberg-Marquardt et de Savitzky-Golay ont été utilisés pour la mise en corrélation des données expérimentales.

```
{DynaFit}
  RandomizationSeed      = 4357 ; | 0 for system time

{ODESolver}
  Iterations             = 1000
  AbsoluteError          = 1.e-14
  RelativeError          = 1.e-8

{Marquardt}
  IterationsPerParameter = 100
  RestartPerturbation    = 0.1
  Restarts               = 2
  RestartsConfidence    = 1
  RobustFit              = n
  EqualizeDatasets      = n
  FixRedundantParameters = y

{ConfidenceIntervals}
  LevelPercent          = 95
  OnlyConstants         = n
  InferenceBand        = confidence ; | prediction
  JointProbability      = n
  MaxSteps              = 30
  SquaresIncreasePercent = 0 ; | 10 for continuous assays

{DifferentialEvolution}
  PopulationSizeFixed    = 0
  PopulationSizeMinimal  = 1000
  PopulationSizePerParameter = 5
  PopulationSizePerOrderOfMag = 3
  MinimumGenerationsPerParameter = 5
  MaximumGenerationsPerParameter = 100
  MaximumEvolutions     = 6
  MinimumEvolutions     = 6
  RandomSeed            = 1234

{Constraints}
  Constants              = 1000000
  Responses              = 1000000
  Concentrations        = 1000
  AllParametersConstrained = y
  AllParametersRelativeBound = 1000000

{Filter}
  PointsPerDataset      = 0
  ExponentialSpacing    = n
  ReadEveryNthPoint    = 0
  SkipFirstNPoints     = 0
  TimeMin               = 0
  TimeMax               = 0
  TimeShift             = 0
  TimeFirstMesh        = 0
  TimeInitialRate      = 1
  PrintInitialRate     = y
```

SmoothData = n
 SmoothingMethod = savitzky-golay ; | average
 SavitzkyGolayWindow = 10
 SavitzkyGolayDegree = 4
 ExtrapolationMethod = quadratic ; | linear
 SmoothingPasses = 4
 AverageReplicates = n
 ZeroBaselineSignal = n

{PieceWiseLinearFit}

Points = 0
 Segments = 4
 Time = 0
 Overlap = n

{Output}

UseDefaultDirectory = y
 Autocorrelations = n
 InferenceBands = n
 WriteTXT = n
 WriteEPS = n
 ColorEPS = y
 ResidualsEPS = y
 WriteTeX = n
 XAxisLabel =
 YAxisLabel =
 BlackBackground = y
 IncludeYZero = n
 InitialRateDigits = 4
 StartDefaultBrowser = n
 PlotRatesLogarithmic = n
 PlotStateLogarithmic = n

{MonteCarlo}

PerformInitialFit = y
 Runs = 1000
 RandomizationMethod = simulate ; | shuffle | shift
 Distribution = normal ; | cauchy |

logistic | uniform

StandardDeviationSource = fit ; | data | explicit
 StandardDeviation = 1.2
 SignificantDigits = 4
 HistogramBuckets = 20
 TruncateMeanPercent = 5
 ColorOutput = y
 RandomizationSeed = 1267
 ConcentrationErrorPercent = 0
 OriginalEstimates = n
 ConfidenceLevel = 100

{EstimateScan}

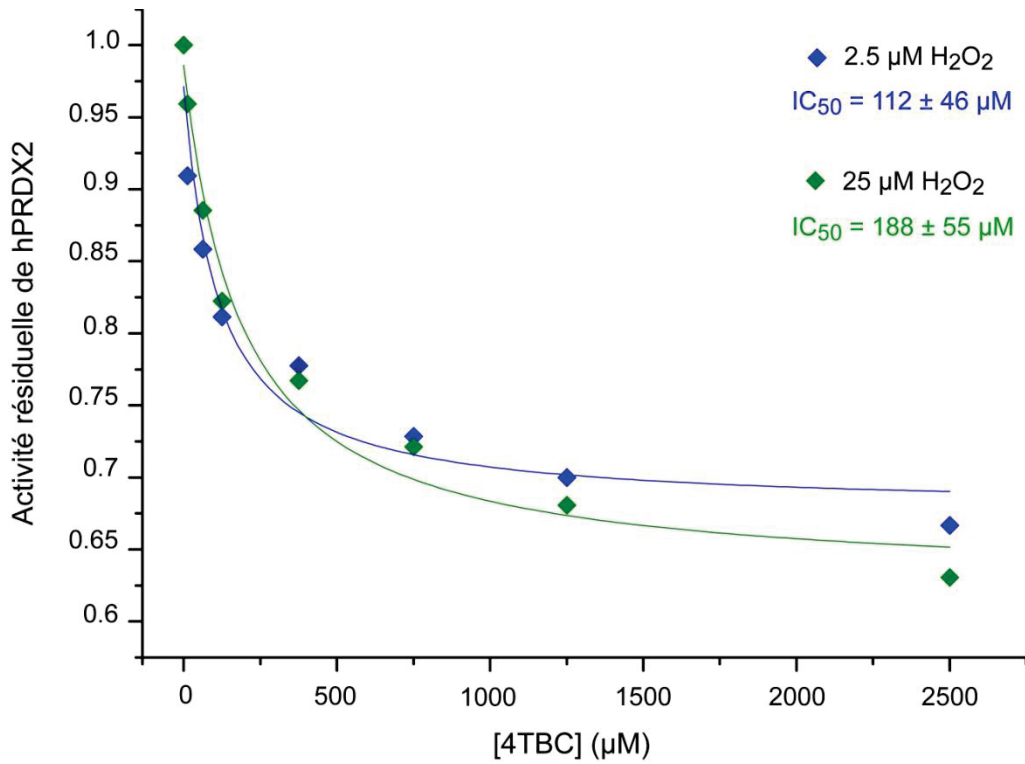
ReportSizeMax = 1000
 RefineMarquardt = 10

{ExponentialFit}
Degree = 4
Automatic = y
AllowOscillations = n
TinyAmplitudes = 0 ; or 0.000001
RefineMarquardt = n

{OptimalDesign}
Algorithm = AS ; | DE | BFGS
Function = D ; | T | E | V ; see
manual

Annexe 7. Comparaison des valeurs d'IC₅₀ obtenus pour [H₂O₂] = K_M et [H₂O₂] = 10K_M

Expérience effectuée sur le système hPRDX2 - 4TBC. IC₅₀ considérés comme non significativement différents à l'erreur près sur la valeur du fit.



PUBLICATIONS

Funnel-Metadynamics and Solution NMR to Estimate Protein–Ligand Affinities

Laura Troussicot,[§] Florence Guillière,[§] Vittorio Limongelli,^{¶,#} Olivier Walker,[§] and Jean-Marc Lancelin^{*,§}

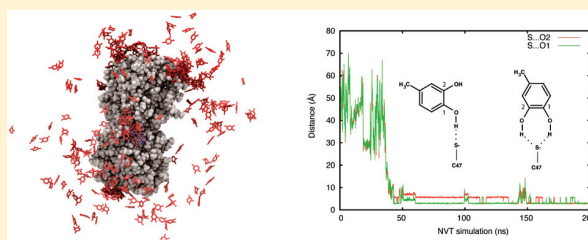
[§]Institut des Sciences Analytiques, UMR 5280, CNRS, Université de Lyon, Université Lyon 1, ENS Lyon -S, rue de la Doua, F-69100 Villeurbanne, France

[¶]Faculty of Informatics, Institute of Computational Science, Università della Svizzera Italiana (USI), via G. Buffi 13, CH-6900 Lugano, Switzerland

[#]Department of Pharmacy, University of Naples “Federico II”, via D. Montesano 49, I-80131 Naples, Italy

Supporting Information

ABSTRACT: One of the intrinsic properties of proteins is their capacity to interact selectively with other molecules in their environment, inducing many chemical equilibria each differentiated by the mutual affinities of the components. A comprehensive understanding of these molecular binding processes at atomistic resolution requires formally the complete description of the system dynamics and statistics at the relevant time scales. While solution NMR observables are averaged over different time scales, from picosecond to second, recent new molecular dynamics protocols accelerated considerably the simulation time of realistic model systems. Based on known ligands recently discovered either by crystallography or NMR for the human peroxiredoxin 5, their affinities were for the first time accurately evaluated at atomistic resolution comparing absolute binding free-energy estimated by funnel-metadynamics simulations and solution NMR experiments. In particular, free-energy calculations are demonstrated to discriminate two closely related ligands as pyrocatechol and 4-methylpyrocatechol separated just by 1 kcal/mol in aqueous solution. The results provide a new experimental and theoretical basis for the estimation of ligand–protein affinities.



Fundamental properties of proteins include their capabilities to interact selectively with other molecules for particular functions as enzyme catalysis or specific receptors signaling, for instance.¹ The structural dynamics of proteins are of prime importance for their interacting properties, and the conformational changes induced upon the binding of a ligand may be the basis of their biological regulations widely referred as allostery.² Since the mid 20th century, crystal structures of proteins were mostly used to describe their chemical architecture both in their free and in their different bound forms to their ligands in the solid crystal state.³ For soluble proteins, NMR introduced the possibility to study proteins in equilibrium between their different states in solution *in vitro*⁴ and even *in vivo*.⁵ For fast exchanges at the NMR time scales most of the NMR observables (isotropic/anisotropic chemical shifts, *J*-couplings, NOEs, etc.) are time-averaged and weighted over the ratio of the exchanging populations.⁶ Currently, the most popular molecular modeling protocols widely used to interpret the results or to predict the binding include docking algorithms which are oversimplified.^{7,8} These popular algorithms do not consider the dynamic properties of proteins beyond the amino acid side chain, ignore the explicit solvent and salts in solution, and clearly lead to severe limitations and possible false or elusive conclusions.

More recently, molecular dynamics (MD) simulations, using still empirical but experienced molecular force fields, conducted in an explicit solvent cell have demonstrated to describe successfully the evolution of the system under investigation in hundreds of nanoseconds (ms) up to the microsecond (μ s) time scale using massively parallel calculations on equilibrated molecular systems of more than 50,000 atoms.^{9,10} These simulations reach the time scales usually detected by the common NMR observables such as chemical shift, whose perturbation can be used to estimate the protein affinity to their ligands.¹¹ A further advance in this field is represented by the use of enhanced sampling techniques such as metadynamics and its derived form funnel-metadynamics (FM),^{12–14} that have been successful to calculate the absolute protein–ligand binding free-energy with excellent convergence criteria, while keeping track of the whole binding-unbinding process. This offers a set of completely new tools for modeling the interaction equilibrium in solution that can contribute to a bright future in the area of medicinal chemistry design of protein ligands and even in the engineering of protein interactions. Here we show for the first time, with the protein human peroxiredoxin 5 (PRX5), that a computational protocol based on FM simulations correlates

Received: November 4, 2014

Published: December 30, 2014

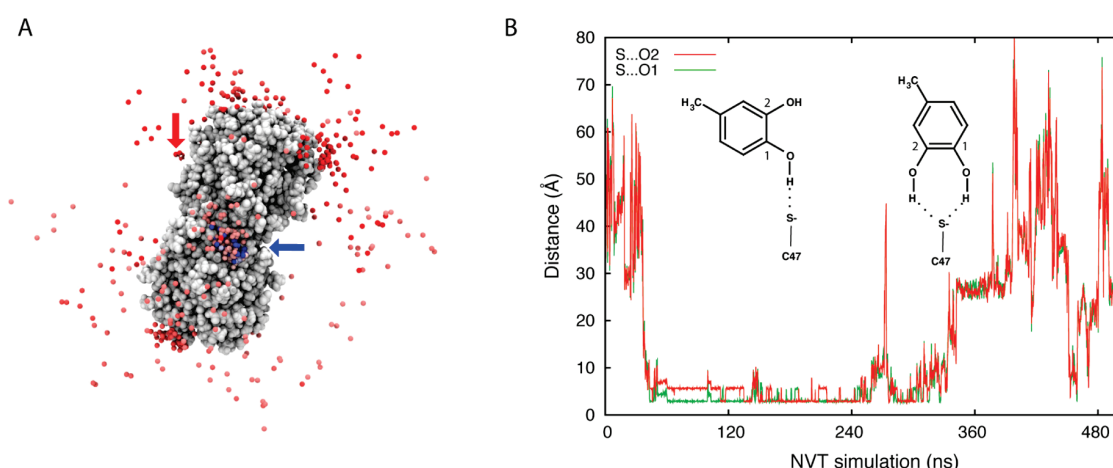


Figure 1. Unbiased NVT MD of human PRX5 and 4-methylcatechol in explicit saline (150 mM NaCl) aqueous solution where a specific binding to PRX5 is observed from a starting unbound state of 4-methylcatechol. (A) PRX5 CPK model and ball representation of O1 atom of 4-methylcatechol plotted every 100 ps steps of the first 200 ns of a 500 ns NVT production MD. The starting ligand position is indicated with a red arrow, while the final positions in the active site of the PRX chain A with a blue arrow. A color continuum from red to blue via white is applied along the MD steps for O1 atom. (B) O1 (green trace) and O2 (red trace) distances from the chain A catalytic thiolate S of C47 as the function of MD steps. 4-methylcatechol docks to the active around 40 ns with characteristic H-bonding patterns with one (different S to O1/O2 distances) or two H-bonds (similar S to O1/O2 distances).

closely to the solution NMR observables used to study the PRX5 selective affinity for simple molecules of the catechol derivatives. The catechol derivatives, pyrocatechol and 4-methylpyrocatechol, were recently found to selectively bind to the active site of PRX5 with a low-affinity. PRX5 is a member of the peroxiredoxin family involved in the peroxide homeostasis and peroxide signaling involved in crucial cell functions.^{15,16} PRXs catalyze the decomposition of hydroperoxides via a complex catalytic pathway involving other proteins. The shape of the active sites of PRXs varies among the PRX family members.¹⁷ The active site consists of central active thiolate and is thought to accept different organic peroxides along with hydrogen hydroperoxide. Our results indicate that the PRX5 affinity to their ligands can be predicted through computations with enough precision regarding the experimental NMR data in solution. The FM results confirm and complement the experimental findings, providing an extensive background information for the protein–ligand binding–unbinding process. This study provides a new way to describe the affinity equilibrium of interacting molecules in solution.

RESULTS

Unbiased MD of PRX at High Concentration of Ligands. To gain a first analysis of the protein–ligand binding and to obtain the collective variables necessary for metadynamics, two different protocols were used. For the first protocol, the homodimer was equilibrated without ligands in the active sites. Whereas the second protocol, oxidized dithiothreitol (DTT) was left in the crystal complex in one of the two active sites (chain A) of the homodimer, while the second active site was free of ligands. In both MD simulations, we positioned 6 molecules of ligands (total concentration in the water box of ~ 15 mM) in the solvent near the two active sites (chain A and chain B) at distances ranging from 15 to 25 Å (measured as the distance between the thiolate SG atom of active site Cys and one of the two oxygen atoms of ligands). A production run of 500 ns was performed to analyze the binding and unbinding events of the ligands within the protein.

In the second MD simulation, DTT unbinds in the first 10 ns from the occupied active site. As illustrated in Figure 1, one of the 6 ligands binds to the liberated active site of the chain A after ~ 40 ns of free diffusion in the solvent and at the surface of the protein. The ligand binds to the active site with a RMS fluctuation of ~ 1 Å for until 260 ns, then it unbinds shortly before rebinding and diffusing again in the solvent and the protein surface. The trajectories of the other ligands are given in Figure S1.

Several binding events occurred on the other active site of chain B. When the ligand is bound, two different binding modes are exchanged during the MD. The major one is strictly comparable to the positions that were described in all crystal structures of PRX for different ligands (benzoate, oxidized DTT, and catechols, see Figure S2).^{18–21} This conformation shows a H-bond formed by OH at position 1 of 4-methylcatechol with an angle between the thiolate SG sulfur of Cys47, O1 and O2 atom of the ligands (S–O–O angle) of about 180° proposed to mimic the reaction transition state.²⁰ This conformer exchanges with another conformation where two H-bonds are formed with the two hydroxyls of the ligand, as shown in Figure 1, with a S–O–O angle of ca. 60° . In the other free active site of the homodimer, a ligand initially close to the active site binds quickly in about 5 ns, first with the double H-bonded conformation (S–O–O angle 60°), then changing into the other single H-bonded conformation (S–O–O angle 180°). These two binding conformers are also correlated with the monitoring of a protein–ligand torsional angle as Cys47SG \cdots O1–C1–C2 (ligand). The conformers with the S–O–O angle 180° has a torsion SG \cdots O1–C1–C2 of $\pm 180^\circ$ and those with a S–O–O angle around 60° has a torsion SG \cdots O1–C1–C2 around 0° (see Figure S2). During the 500 ns MD run, several unbinding periods are observed ranging from a few ns to 20 ns, leaving the active site free. A direct exchange between two catechol ligands was also observed during a short period of about 2 ns before exchanging back to the first ligand.

Free-Energy Calculations Using FM Simulations. To provide more insight into the thermodynamics ruling the

binding of the ligands to PRX, well-tempered metadynamics simulations were performed.^{22,23} This enhanced sampling technique allows the simulation of long time scale events, such as ligand/protein binding, in a reasonable computational time duration. At the end of the calculation, the free energy landscape of the investigated process can be computed using the history-dependent bias potential added during the simulation on few degrees of freedom of the system, called collective variables (CVs). This technique has been successfully used to study complex and long time scale biological processes like DNA folding²⁴ and ligand–protein and ligand–DNA binding.^{25–27} In the present study, in order to investigate the ligand–PRX binding process and calculate the ligand binding free energy, a recently developed metadynamics-based approach, called funnel-metadynamics (FM), was used.¹⁴ These simulations use a funnel-restrained potential that reduces the conformational space to explore the unbound state, thus increasing the number of binding events observed during the sampling. In such a way, FM calculations lead to a quantitatively well-characterized binding free-energy surface (FES). It is important to stress that when the ligand is inside the funnel, no external potential is applied, and the system works under the standard metadynamics regime. Thus, the shape of the funnel restraint potential was purposely chosen to include the PRX binding site within the cone section (Figure 2).

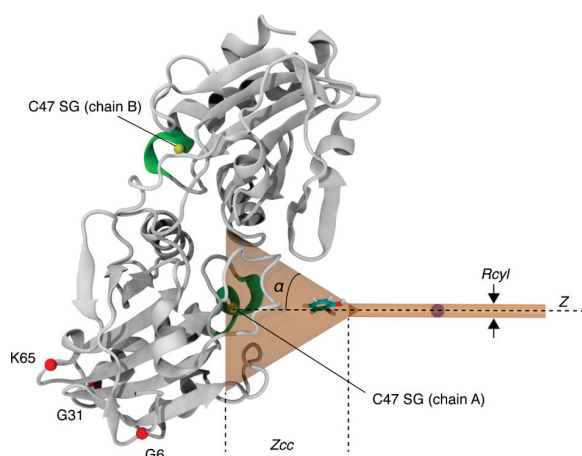


Figure 2. Funnel restraint potential applied to PRX5 enzymes with catechol ligand. The catalytic thiolate of C47 in the α -helix colored in green is represented with a yellow sphere on chains A and B of the PRX5 homodimer. α of G6, G31, and K65 of chain A used as positional restrains to inhibit the whole protein diffusion is represented with red spheres. The cylinder part is aligned and centered on the Z axis with a radius $R_{\text{cyl}} = 1 \text{ \AA}$. The funnel potential is limited to 35 \AA from the Z-axis origin (SG atom of Cys47) by a wall potential represented by a violet sphere. The cone region is defined by a vertex height Z_{cc} of 18 \AA from the origin and an angle α of 1.1 rad.

This can be done by properly setting the angle α and the distance z_{cc} parameters of the funnel restraint potential (see Materials and Methods section and ref 14 for details). In such a way the external potential does not influence the ligand exploration of the binding site, $z < z_{\text{cc}}$, while a cylindrical restraint potential is applied when the ligand is in the unbound state, $z > z_{\text{cc}}$.

Using this protocol the ligand–protein binding constant, K_{b} , can be computed as follows:

$$K_{\text{b}} = \pi R_{\text{cyl}}^2 \int_{\text{site}} dz e^{-\beta[W(z) - W_{\text{ref}}]} \quad (1)$$

where πR_{cyl}^2 is the surface of the cylinder used as restraint potential, while the potential $W(z)$ and its value in the unbound state, W_{ref} , can be derived from the potential of mean force (PMF) obtained through metadynamics calculations. β is constant, where $\beta = (k_{\text{B}} T)^{-1}$, k_{B} is the Boltzmann constant and T the temperature of the system.

The equilibrium binding constant, K_{b} , is directly related to the absolute protein–ligand binding free energy, ΔG_{b}^0 , through the following formula:

$$\Delta G_{\text{b}}^0 = -\frac{1}{\beta} \ln(C^0 K_{\text{b}}) \quad (2)$$

where $C^0 = 1/1660 \text{ \AA}^{-3}$ represents the standard concentration 1 M.

Note that the equilibrium binding constant, K_{b} , is the inverse of the dissociation equilibrium constant of the ligand–protein complex, K_{D} , usually measured by NMR experiments (see High-Precision Dissociation Constant Determination Using Solution NMR section).

Using FM simulations, Limongelli et al. have studied two different protein–ligand systems, namely trypsin/benzamidine and cyclooxygenase II (COX II) in complex with a potent inhibitor.¹⁴ These authors have not only described with atomic details the complex ligand–protein binding mechanism but also computed the ΔG_{b}^0 values for both systems. The calculated ΔG_{b}^0 and the derived K_{D} fall in the range of the experimentally determined values.

The active site of PRX5 is directly accessible from the solvent, and the cone part contains all possible interacting residues of PRX5. FM simulations were then performed for more than 500 ns applying the metadynamics bias on the distance CV defined as the projection on the Z axis of the distance between Cys47 SG and the center of mass of the ligand heavy atoms. The statistics of two further CVs, specifically the distance from the funnel Z axis of the ligand center of mass and the protein–ligand torsion Cys47 SG...O1–C1 (ligand) or the SG...O1–O1 angle (see above), were collected along the FM simulations to evaluate the conformational space explored by the system (Figure S3). During the FM simulations the ligands exchange several times between the different bound conformational states diffusing in the solvent. Binding and unbinding events occurred 8–10 times over the whole 500 ns of FM. The binding events take approximately 10–30 ns, separated by unbound periods of 20–30 ns. The calculation convergence was monitored using the bias deposition as well as the potential mean force (PMF) $W(z)$ and ΔG_{b}^0 calculated as described before.¹⁴ FES as a function of CVs different from that biased in the FM calculation (Figure 3) can be reconstructed using the potential added to the system and a reweighting algorithm.²⁸

Absolute Binding Free-Energy. To have a quantitatively well-characterized free-energy profile, a number of recrossing events between the different states visited by the system should be observed.¹⁴ As shown in Figure S3 both systems visit several times the bound ($z < 8 \text{ \AA}$) and the unbound states ($z > 20 \text{ \AA}$) during the simulation. Using $R_{\text{cyl}} = 1 \text{ \AA}$ and $W(z)$ obtained from metadynamics (Figure 3 upper-left panel) in eqs 1 and 2, the estimate of ΔG_{b}^0 for catechol and 4-methylcatechol converges to -3.0 ± 0.2 and -4.2 ± 0.3 kcal/mol, respectively. To provide a picture of the convergence of the binding free-energy estimation, the free-energy difference between bound and unbound states has been computed as a function of the simulation time for either system (Figure 3 lower-left panel).

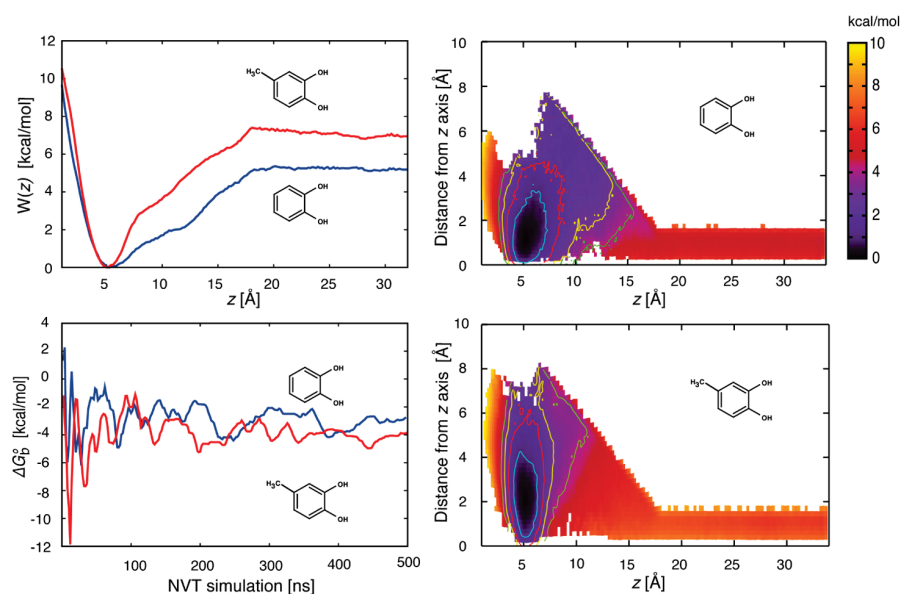


Figure 3. PMF and FES represented as a function of the different collective variables used in FM calculations and evolution of the absolute binding free energy (lower-left panel). Upper-left: compared PMF $W(z)$ of catechol (blue trace) and 4-methylcatechol (red trace) as a function of the Z-axis projection distance between Cys47 SG and the center of mass of the ligand heavy atoms. Upper/lower-right: FESs of the PRX5-ligand system for catechol (upper) and 4-methylcatechol (lower) as a function of the protein–ligand distance projection on Z axis and the protein–ligand distance from Z axis. The FESs show the lowest energy basins found by the FM calculations. Lower-left panel: evolution of absolute binding free energy during 500 ns of FM simulation for catechol (blue trace) and 4-methylcatechol (red trace). Using PMF $W(z)$ and $R_{\text{cyl}} = 1 \text{ \AA}$ in eqs 1 and 2, the estimate of ΔG_b^0 for catechol and 4-methylcatechol converges to -3.0 ± 0.2 and -4.2 ± 0.3 kcal/mol, respectively. The uncertainty is calculated as the SD from the asymptotic value of the absolute protein–ligand binding free energy obtained in the last part of the simulation.

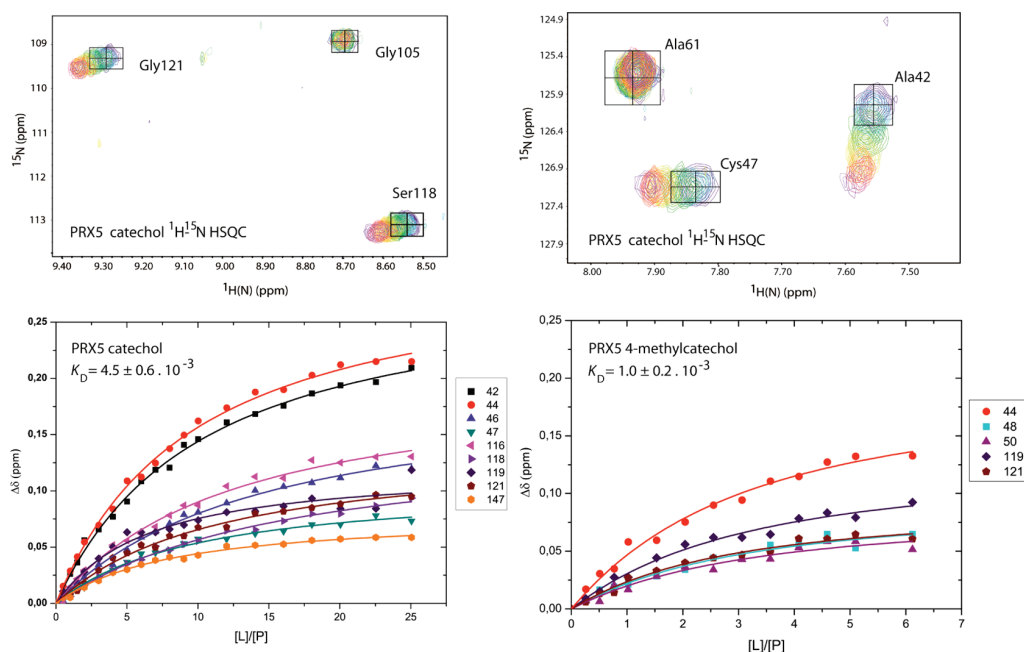


Figure 4. High-precision determination of dissociation constants K_D of human PRX5 bound to catechol and 4-methylcatechol using solution NMR. Upper-panels: parts of the overlaid ^1H – ^{15}N HSQC spectra of PRX5 at various concentrations of catechol in aqueous solution, 28°C and pH 7.4. Chemical shift indicated with boxed crosses are the initial NMR chemical shift with no ligand. Lower two panels: plots of the NMR chemical shift perturbation ($\Delta\delta$) as a function of ligand-to-protein concentration ratio ($[L]/[P]$) for the amino acids sequence position indicated on the graphs. Data were nonlinearly fitted to the 1:1 protein–ligand model $P + L \rightleftharpoons PL$ with a single dissociation K_D .

Furthermore, using eq 1 and considering that K_b is the inverse of K_D , we could estimate K_D for catechol and 4-methylcatechol, which is $6.9 \pm 2.1 \times 10^{-3}$ and $0.9 \pm 0.4 \times 10^{-3}$, respectively.

High-Precision Dissociation Constant Determination Using Solution NMR. NMR is ideally suited for the analysis of protein interactions with dissociation constants in the μM

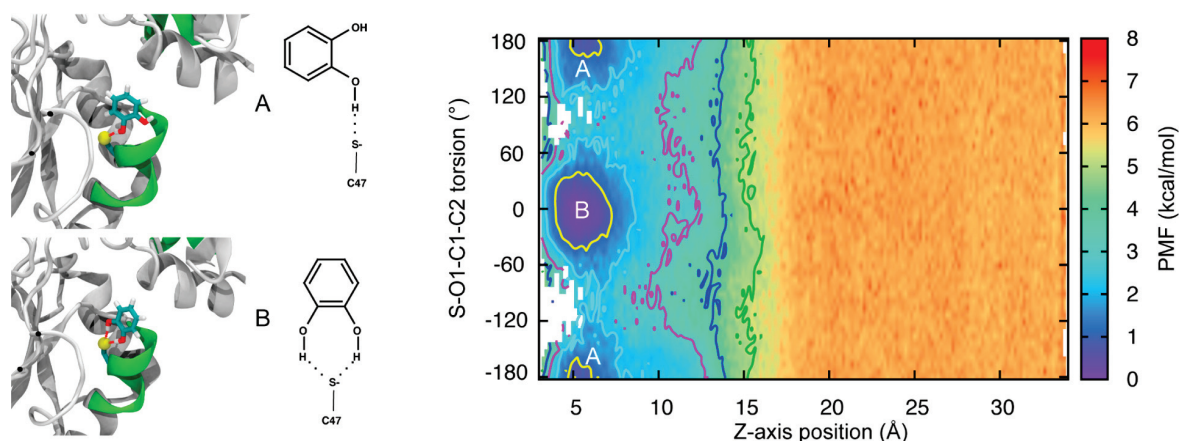


Figure 5. FES of catechol interacting basins with PRX5 evaluated using FM. Free energies are represented using iso-energetic contours spaced by 1 kcal/mol from the minima at 0 and 5 kcal/mol and a color continuum from red (highest free energy) to blue (lowest free energy) as a function of the positional projection of catechol center of mass along the Z axis defined according to the Material and Methods section, and the torsion defined by the catalytic thiolate SG atom of PRX5 Cys27 residue and the O1, C1, and C2 atoms of catechol. Map minimum B represents the double H-bonds binding mode, while energetic A basin represents the single aligned H-bond similar to the crystal ligand positions (see the text).

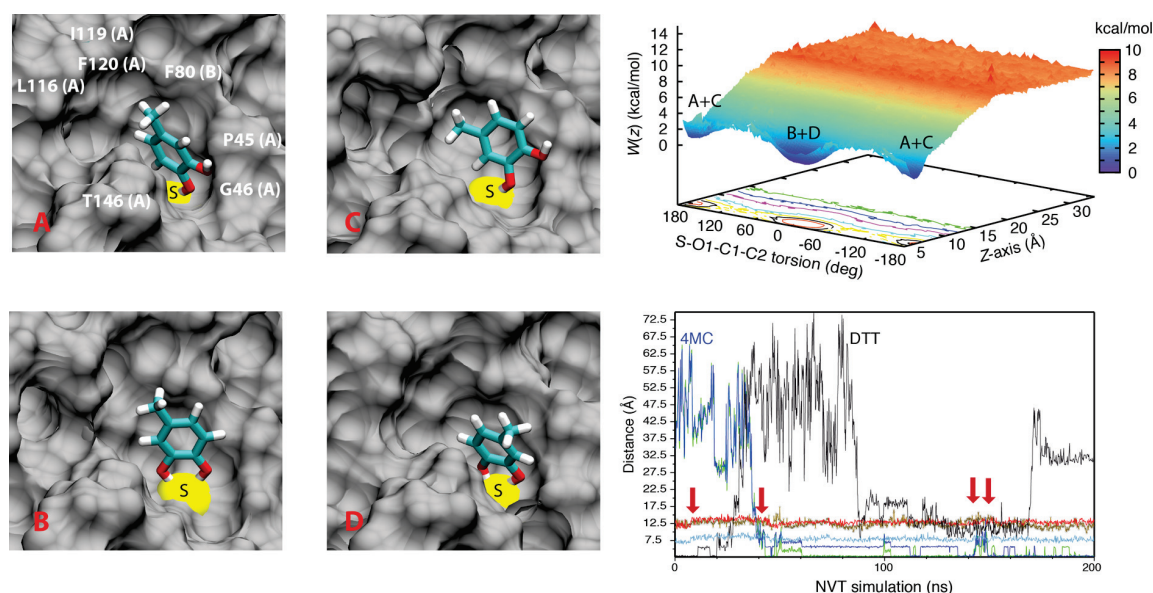


Figure 6. Binding FES of the human PRX5/4-methylcatechol complex obtained through FM (upper right), NVT unrestrained dynamics trajectories (lower right), and representation of the lowest energy ligand binding modes (left). Snapshots represent the surface accessible to the solvent of the PRX5 active site in gray and the SG atom of catalytic Cys47 in yellow. 4-methylcatechol is represented with sticks, and the surrounding amino acids are indicated either from the chain A or B of the PRX5 homodimer. Right diagrams, top: FES representation as a function of two collective variables used in funnel-metadynamics. Z-axis projection distance between Cys47 SG and the center of mass of the ligand heavy atoms and the protein–ligand torsion angle defined by PRX5 Cys47 SG, 4-methylcatechol O1, C1, and C2 atoms. Lower-right diagram: interatomic distances as a function of unrestrained NVT dynamics time steps during the first 200 ns of the 500 ns run. Protein (Cys47)-ligand distances measured for 4-methylcatechol (4MC) and DTT are plotted in green (SG-O1), blue (SG-O2), and black (SG to the DTT oxygen H-bonded to Cys47 in the X-ray structure), respectively, indicating the binding/unbinding events along the time steps (red arrows). Interatomic distances within the PRX5 active site including Cys47 SG with Leu116 CG (brown), Ile119 CB (red) and Phe120 (light blue) are graphed.

to mM range in the fast exchange equilibrium on the NMR time scale.^{29,30} ^1H – ^{15}N NMR data of PRX5 can be used to titrate the reduced form PRX5 against catechol and 4-methylcatechol.³¹ Since the oxidized form dithiothreitol (DTT), a reagent commonly used to reduce disulfide bridges of proteins, was reported as a ligand of PRX5,²⁰ we preferred to use the unrelated reducing agent tris(2-carboxyethyl)phosphine (TCEP) hydrochloride, as not to interfere with the ligand binding. Figure 4 shows overlay of ^1H – ^{15}N HSQC spectra recorded after each increasing

concentrations of the ligands. Specific amino acids are affected by the interaction and present significant chemical shift perturbations (CSPs). The corresponding CSPs were then quantified carefully. The most pronounced CSPs are the ^1H – ^{15}N amide group of amino acids in the active site, near the catalytic cysteine (Cys47). Residues Ala42, Thr44, Gly46, and Cys47 for catechol and Thr44, Ser48, and Thr50 for the methylated homologue are mostly affected. Some amide groups of amino acids located in the loop and the α -helix adjacent to the active site also present

significant CSPs (Leu116, Ser118, Ile119, Gly121, and Thr147 for catechol and Ile119, Gly121 for 4-methylcatechol). Observed significant CSPs suggest that interaction of PRX5 with catechol and 4-methylcatechol is indeed specific despite their low-affinity.

When the most significant CSPs are plotted against the ratio ligand-to-protein concentrations $[L]/[P]$, hyperbolic isotherms are obtained as shown in Figure 4. The data can be best fitted using a nonlinear fitting protocols corresponding to a 1:1 protein–ligand binding model with a single K_D . Overall catechol induced more CSPs than 4-methylcatechol. Catechol induced to the most (Ala42 and Thr44) 0.2 ppm of CSPs, while 4-methylcatechol induced to the most CSPs between 0.1–0.15 ppm. However, the averages of the data (each titration was duplicated) shown in Figure 4 lead to K_D 's of 4.5 ± 0.6 and 1.0 ± 0.2 kcal/mol at 301 K in phosphate buffer saline (PBS) solution pH 7.4 (NaPi 10 mM, NaCl 137 mM, KCl 3.3 mM) TCEP 2 mM, for catechol and 4-methylcatechol, respectively.

DISCUSSION

Details of the Ligand–Protein Interaction. When the ligand occupies the protein active site, the two binding modes (protein–ligand S–O–O angles 60° and 180° , and protein–ligand SG–O1–C1–C2 torsion 0° and 180° , respectively) exchange regularly, and the ligand samples the whole conformational space dynamically. For catechol, which owns a C_{2v} symmetry, the two SG–O1–C1–C2 torsion values correspond to two different binding conformations, as shown in Figure 5. For 4-methylcatechol with a C_i symmetry the two SG–O1–C1–C2 torsion values correspond to more possible conformations in which the methyl group could interact differently within the PRX5 active site. In the crystal state, the methyl group is oriented in the active site interacting with neighboring hydrophobic side chains such as Leu116, Ile119, Phe120, and Phe80 (Figure 6A). Due to the protein dynamics, the active site changes its shape during the ligand binding and unbinding events. This is clearly shown in Figure 6 where the distances between Cys47 SG atom, and the side chains of Leu116 and Ile119 are plotted during a standard MD calculation. In particular, one can observe that, on average, these distances elongate during the unbinding events and shorten in the binding by about 2 Å.

The different ligand binding conformations could not be discriminated just by observing the FES associated with the two protein–ligand distance CVs given in Figure 3 (right panels).

Catechol. The different ligand binding modes can be identified reconstructing the FES as a function of the distance CV on Z axis and the protein–ligand SG–O1–C1–C2 torsion (Figure 5). This operation is possible using the reweighting algorithm of Bonomi et al.²⁸ Looking at the FES shown in Figure 5 two free-energy minima can be found. The first, basin A, corresponds to the ligand binding mode with a single protein–ligand H-bond, including the conformations found by X-ray.^{20,21} The second basin, B, is energetically equivalent to A and represents the binding mode with two protein–ligand H-bonds.

4-Methylcatechol. As for catechol, all the bound 4-methylcatechol conformations fall in the free-energy basins reported in Figure 6. Here, two most populated binding modes are found: conformer A, which is similar to the crystal conformation with a single H-bond, and conformer B, which shows the chelate H-bond interaction. In this state, the methyl group is orientated to the hydrophobic cavity formed by

residues such as Phe80 (chain B), Phe120, and Leu116. Two less populated states are also present in both energy basins. Conformer C is a minor state of basin A where the ligand is slightly moved from A due to the steric hindrance of the ligand methyl group. Conformer D is a minor state of basin B where the methyl group points toward the solvent where it is not possible to form favorable hydrophobic interactions. The weaker ligand–protein interactions formed in C and D if compared with A and B lead us to consider these poses the first binding event of 4-methylcatechol in the active site before reaching its final position. Alternatively, they can be considered the first unbinding event of the ligand from the catalytic site of the enzyme.

Calculated Absolute Binding Free Energies and Experimental High-Precision Dissociation Constants Determined by NMR. Table 1 summarizes the calculated

Table 1. Comparison of Calculated and Experimental Dissociation Constants K_D and Absolute Binding Free Energy ΔG_b^0 of Catechol and 4-Methylcatechol for Human PRX5

	catechol	4-methylcatechol
ΔG_b^0 ^a		
FM ^b	-3.0 ± 0.2	-4.2 ± 0.3
NMR	-3.2 ± 0.1	-4.1 ± 0.1
K_D (10^{-3})		
FM ^b	6.9 ± 2.1	0.9 ± 0.4
NMR	4.5 ± 0.6	1.0 ± 0.2

^akcal/mol. ^bMean value over the last 100 ns of 500 ns FM.

absolute binding free energies ΔG_b^0 using FM protocol and the dissociation constants K_D determined using NMR. For each ligand the experimental value matches very well the calculated one.

Since both ligands are soluble in water, we checked that the presence of 0–8% DMSO does not induce a significant difference. DMSO alone causes some limited CSP of ^1H – ^{15}N other amino acids including Ser48 and Gly46 that were also used for the K_D determinations affecting marginally the determined dissociation constant values. Previously published data, determined using NMR saturation transfer difference spectroscopy, for the catechol-PRX5 system reported a K_D of $3.3 \pm 0.5 \times 10^{-3}$ that is very close to our estimate.³¹ The use of DTT instead of TCEP reducing agent does not affect significantly the measured values. We just noticed that DTT caused a pronounced long-term precipitation of the PRX, modifying the PRX5 concentration. However, if the NMR titration is done within a day or less, no significant differences of the data were noticed.

Structural Source of the Difference in ΔG_b^0 between the Two Ligands of PRX5. Based on the analysis of the FESs, the source of stabilizing interactions between the catechol ligands and PRX5 is clearly the possibility to form specific patterns of H-bonds between its two hydroxyl groups and the thiolate of the catalytic Cys47.

Given the estimated pKa of catalytic PRX Cys ranging from 5 to 6, more than 90% of Cys47 is a thiolate in solution at pH 7 and the source of the reactivity of PRX's toward their substrate peroxides.²⁰ The specific conformer's geometry of the type A in Figure 5 was proposed to be an analog of the transition state²⁰ and lead to a well-defined free-energy basin. This sort of stabilizing interaction is reproduced with 4-methylcatechol

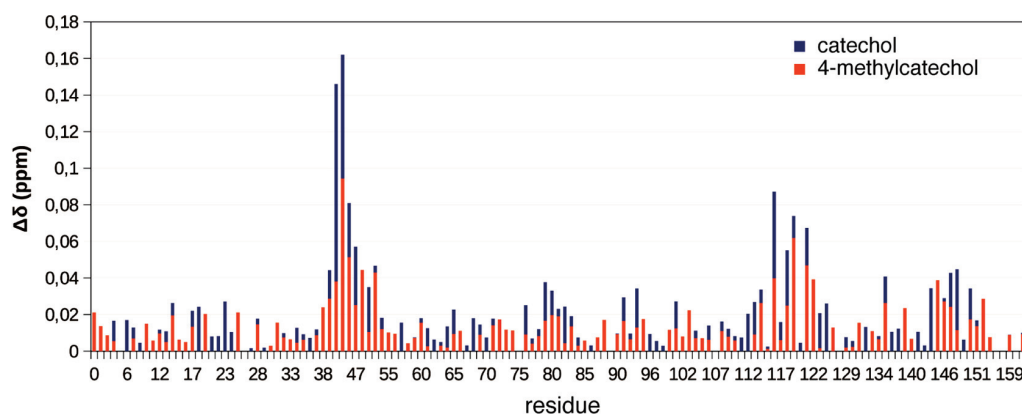


Figure 7. Superimposed histograms of the chemical shift perturbations $\Delta\delta$ caused to human PRX5 ^1H – ^{15}N NMR signals by catechol (blue bars) and 4-methylcatechol (red bars) at concentrations twice their respective determined dissociation constant K_D 's giving a comparable 75% of the PRX5 saturation each. Missing data due to spectral overlaps or nonobserved amino acids are not represented.

extending the stabilizing van der Waals interactions laterally with Leu116, Leu119, Phe120 side chains of the chain A and also the side chain of Phe80 of the chain B. Dynamically and on average, the active site contracts upon binding the 4-methylcatechol to optimize those additional van der Waals interactions. As a result, the lowest free-energy basin of this ligand presented in Figure 3 is more extended compared to that of catechol. This contributes, with other factors such as the higher polarity of catechol that leads to more favorable interactions with the solvent, to a gain in ΔG_b^0 of about 1 kcal/mol. We stress that the differences in ΔG_b^0 for the two ligands is accounted for both experimentally using NMR and theoretically by the metadynamics calculations.

MD Calculations and NMR. Our work demonstrated for the first time that free-energy calculations using enhanced sampling simulations correlate with NMR-determined affinities for weakly interacting protein–ligand system with dissociation constants at room temperature in the order of 10^{-3} . Recent studies with MD and FM were conducted for more affine systems with K_D in the range from 10^{-6} to 10^{-9} .^{10,14} NMR observables and isotropic chemical shift in particular are essentially an average over the NMR time scales (ms to s for chemical shifts) of the conformations sampled by the system. The relevant modeling of the NMR observables should then rely on the full statistical description of the complete exchanging conformations.

The use of bias in metadynamics and external potentials as the funnel-shaped potential used in FM protocols allows faster and more complete sampling than standard simulations. The two essential states of an interacting system (the free state in the solvent and the bound forms) exchange more, and the several possible bound conformers of low energies are better sampled. At the convergence, ΔG_b^0 reaches a mean stable value, and the system motion can be considered diffusive. Thus, the use of biased MD simulation allows a quantitative description of the interactions. This description is by far more accurate than that obtained using faster but less accurate approaches, such as molecular docking, where the intrinsic dynamical character of proteins is neglected. An example of the relevance of using MD-based protocols in modeling NMR parameter can be given by the consideration of the two ligands CSPs induced at a comparable ratio between free protein and bound protein. For instance, at ligand concentrations near the determined K_D (i.e., at 50% of the saturation) as shown in Figure 7,

4-methylcatechol induced on average significantly less CSPs than catechol despite a lower K_D . The difference in CSPs can be ascribed to the larger number of ligand–protein interactions formed by 4-methylcatechol if compared with those of catechol. This aspect is explained by the FESs computed for both ligands. In fact, while catechol shows a certain free-energy basin, 4-methylcatechol has a free-energy basin extending to a more larger distance from the Z axis (vertical axis in Figure 3). This reflects a greater number of iso-energy binding conformations. All the protein motions due to these interactions are averaged in the time scale of the NMR experiment, resulting in lower CSPs for 4-methylcatechol. Such considerations, and the possibility to describe quantitatively through simulations the whole exchanging systems at a time scale relevant to NMR spectroscopy, were nearly ignored until now for interacting protein–ligand systems.

We believe that this first approach of coupling enhanced sampling calculations to NMR and other microscopic and macroscopic experimental data is the only way to rationally interpret the interacting systems in solution. The resulting knowledge may contribute significantly to the possible design and engineering of the interactions to conceive high-affinity ligand from low-affinity primary hits.

CONCLUSION

We have shown herein that MD, complemented by metadynamics, correlates not only in the energetics of a protein–ligand interacting system (accurate estimates of ΔG_b^0) but also in the whole dynamical exchanges occurring on time scale relevant to NMR spectroscopy. The consequences are important for the comprehensive understanding of the protein–ligand affinity and useful for a proper engineering and design of new interacting molecules.

MATERIALS AND METHODS

Molecular Dynamics. All MD runs were prepared starting from the conformation of the crystal structure of PRX5 complexed with oxidized DTT (3MNG.pdb).²⁰ PRX5 homodimers were constructed from appropriate crystallographic C2 symmetry mates. Simulations were carried out with the AMBER99SB-ILDN force field^{32–34} for the protein and the TIP3P water model for the explicit solvent.³⁵ Crystallographic water molecules were kept to save in particular the low-diffusion water molecules buried in the protein structure. Active site Cys 47 was modeled a thiolate (residue CYM of Amber), and 48 Na^+ and 46 Cl^- were added to account for the experimental conditions

used in NMR for ca. 150 mM NaCl aqueous solution. The Amber charges were applied to protein, ions, and water atoms, and the restrained electrostatic potential charges were used for the ligands using the Antechamber program suite³⁶ and the General Amber Force Field GAFF.³⁷ Using ACEMD code,⁹ the system was minimized and equilibrated under constant pressure and temperature (NPT) conditions at 1 atm and 300 K using a time step of 4 fs thanks to the use of the hydrogen mass repartitioning scheme³⁸ implemented in ACEMD, nonbonded cutoff of 9 Å, rigid bonds and particle-mesh Ewald long-range electrostatics with a grid of 82-83-92 with spacing of 1 Å. The systems were equilibrated first using 500 steps of steepest-descent minimization, followed by running 0.1 ns of the isothermal NVT ensemble, using a Langevin thermostat set at 300 K, followed by 5 ns of the isothermal–isobaric NPT ensemble using a Langevin thermostat at the same temperature and the Berendsen barostat of ACEMD. During minimization and equilibration, the heavy protein and ligands atoms were restrained spacially using 10 kcal mol⁻¹ Å⁻² spring constant. The magnitude of the restraining force constant was then reduced to 1 kcal mol⁻¹ Å⁻² during 100 ps of NVT. The barostat was switched on at 1 atm for 1 ns of NPT simulation. During that period the force constant of the position constrain of all heavy atoms was gradually reduced every 100 ps steps by a factor 0.65 to a final value of about 0.1 kcal mol⁻¹ Å⁻², allowing the systems to relax gently. Finally, the volume was allowed to relax for further 4 ns under NPT conditions reaching a final box size of 82.7-83.7-93.9 Å³. During this run, Cα atoms of the protein and ligands were restrained with a 1 kcal mol⁻¹ Å⁻² harmonic potential to prevent the system reorienting. Then, a production run of 500 ns in the NVT ensemble was performed.

Funnel-Metadynamics.¹⁴ Unrestrained simulations were run at the same temperature and solvent model in the NVT ensemble for the indicated simulation times. The PLUMED plugin²⁸ was used to carry out metadynamics calculations. Since the funnel external potential is fixed in the space,¹⁴ PRX5 dimer is diffusion restrained by three atoms chosen far from the dimer interface and the active site as shown in Figure 2. The bias was added on a distance CV (see Results section). A Gaussian width of 0.35 Å was used, and a Gaussian deposition frequency of 0.5 kcal mol⁻¹ every 2 ps (1 kcal = 4.18 kJ) was initially used and gradually decreased on the basis of the adaptive bias with a ΔT of 3600 K. Trajectories were analyzed using VMD software.³⁹

Protein Production and Purification. Recombinant Human PRDX5 was expressed without its mitochondrial targeting sequence as a N-terminal 6xHis-tagged protein using *Escherichia coli* M15 (pRep4) strain (plasmid coding for the protein was generously given by B. Knoops lab at University of Louvain, Belgium). To produce ¹⁵N uniformly labeled PRDX5, cells were grown at 37 °C in M9 minimal medium containing ¹⁵NH₄Cl 1g/L as sole source of nitrogen and supplemented with thiamine 20 μg/mL, metal traces solution, and the two antibiotics of resistance, kanamycin and ampicillin (50 μg/mL of each). Produced PRDX5 was then purified using nickel-affinity chromatography with a Ni²⁺-NTA column (Qiagen) as previously described³¹ and dialyzed against a PBS pH 7.4 (NaH₂PO₄/Na₂HPO₄ 10 mM, NaCl 137 mM, KCl 3.3 mM). Protein concentration was measured according to the UV absorbance at 280 nm based on a molar extinction coefficient at that wavelength of 5500 M⁻¹ cm⁻¹ calculated⁴⁰ for one chain of PRX5 (concentration expressed in molar concentration of monomer in solution).

NMR Experiments. All spectra were recorded at 28 °C with a Varian Inova 600 MHz NMR spectrometer equipped with a 5 mm standard triple resonance (¹H/¹³C/¹⁵N) inverse probe with a Z-axis field gradient. The concentration of DMSO-*d*₆ did not exceed 8% in the NMR samples, and DMSO's influence was tested separately. All NMR data sets were processed using NMRPipe/NMDraw package,⁴¹ and ¹⁵N chemical shift changes were assigned using NMRViewJ software.⁴² Ligand interactions with U-¹⁵N-PRX5 were characterized using CSPs. A series of ¹H-¹⁵N HSQC experiments were acquired upon addition of specific ligand concentrations. NMR samples in typical series contained 400 μM of the reduced ¹⁵N-labeled protein. TCEP 0.1 M was added to maintain the protein in its reduced form in PBS pH = 7.4 to a final concentration in TCEP of 2 mM.

To avoid possible aggregation of the proteins, an increasing volume of a concentrated stock solution of ligands (110 mM solutions in DMSO or H₂O) was incrementally added until saturation was reached. To derive the corresponding binding constant K_D (see below), spectral perturbations were quantified as the combined amide CSPs Δδ according to the following expression:⁴³

$$\Delta\delta = \sqrt{\frac{\Delta\delta_{\text{H}}^2 + \frac{1}{25}\Delta\delta_{\text{N}}^2}{2}} \quad (3)$$

Dissociation Constant Determination from NMR CSPs.

Titration data were analyzed assuming that the observed chemical shift perturbation Δδ is a weighted average between the two extreme values corresponding to the free (Δδ = 0) and ligand-bound states (Δδ = Δδ_B). PRX5 was assumed to contain two independent binding sites corresponding to the equivalent chains A and B of the symmetrical homodimer. CSPs data were then nonlinearly fitted against the theoretical 1:1 model of protein–ligand binding of the simple binding equilibrium:



Where P stands for the protein, L for the ligand and PL for the protein–ligand complex. This equilibrium has dissociation constant K_D expression:

$$K_D = \frac{[P_0][L_0]}{[PL]} - [P_0] - [L_0] + [PL] \quad (5)$$

where [P₀] and [L₀] are the total concentrations in protein and ligand, respectively, and [PL] the concentration of protein–ligand complex at the equilibrium. A 1:1 model with one ligand molecule bound per PRX5 gives⁴⁴

$$\Delta\delta = \frac{[PL]\Delta\delta_B}{[P_0]} \quad (6)$$

so

$$\Delta\delta = \frac{\Delta\delta_B([P_0] + [L_0] + K_D) - \sqrt{([P_0] + [L_0] + K_D)^2 - 4[P_0][L_0]}}{2[P_0]} \quad (7)$$

where the chemical shift perturbations (Δδ_B) and the dissociation constant (K_D) were fit with nonlinear regression by using an in-house Matlab (The MathWorks, Inc.) based program.⁴⁵ Error margins on final K_D values were calculated using a Monte Carlo simulation assuming 20% of experimental error on protein and ligand concentrations under a cumulative error on volume and weigh-in. All titrations were achieved in duplicate and the results expressed as the averages.

■ ASSOCIATED CONTENT

📄 Supporting Information

Figures S1–S3 providing trajectories collective variable analysis of unrestrained MD and restrained FM calculations. This material is available free of charge via the Internet at <http://pubs.acs.org>.

■ AUTHOR INFORMATION

✉ Corresponding Author

*jean-marc.lancelin@isa-lyon.fr

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We are grateful to Prof. Bernard Knoops and André Clippe from University of Louvain-la-Neuve, Belgium, for the generous gifts of biological materials to express the human peroxiredoxin in *Escherichia coli*. We thank Acellera Company for providing of a complementary 6 month ACEMD multi-GPU software

licence and for the generous gift of a Nvidia Tesla K20 supercomputing unit. Funding of this project was provided by a grant from Région Rhône-Alpes, France, no. 13-0118962-01 to L.T. and J.-M.L., and a grant from the Swiss National Supercomputing Centre (CSCS) under project ID s557 to V.L.

REFERENCES

- (1) Kuriyan, J.; Konforti, B.; Wemmer, D. *The Molecules of Life: Physical and Chemical Principles*; Garland Science: New York, 2012.
- (2) Changeux, J. P. *Protein Sci.* **2011**, *20*, 1119.
- (3) Bernstein, F. C.; Koetzle, T. F.; Williams, G. J.; Meyer, E. F., Jr.; Brice, M. D.; Rodgers, J. R.; Kennard, O.; Shimanouchi, T.; Tasumi, M. *Arch. Biochem. Biophys.* **1978**, *185*, 584.
- (4) Wüthrich, K. *Angew. Chem.* **2003**, *42*, 3340.
- (5) Serber, Z.; Keatinge-Clay, A. T.; Ledwidge, R.; Kelly, A. E.; Miller, S. M.; Dotsch, V. *J. Am. Chem. Soc.* **2001**, *123*, 2446.
- (6) Ernst, R. R. B., G.; Wokaun, A. *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*; Clarendon Press: Oxford, 1987; Vol. 14.
- (7) Goodsell, D. S.; Morris, G. M.; Olson, A. J. *J. Mol. Recognit.* **1996**, *9*, 1.
- (8) Guo, L.; Yan, Z.; Zheng, X.; Hu, L.; Yang, Y.; Wang, J. *J. Mol. Model.* **2014**, *20*, 2251.
- (9) Harvey, M. J.; Giupponi, G.; Fabritiis, G. D. *J. Chem. Theory Comput.* **2009**, *5*, 1632.
- (10) Buch, I.; Giorgino, T.; De Fabritiis, G. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 10184.
- (11) Williamson, M. P. *Prog. Nucl. Magn. Reson. Spectrosc.* **2013**, *73*, 1.
- (12) Limongelli, V.; Bonomi, M.; Marinelli, L.; Gervasio, F. L.; Cavalli, A.; Novellino, E.; Parrinello, M. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 5411.
- (13) Barducci, A.; Bonomi, M.; Parrinello, M. *WIREs Comput. Mol. Sci.* **2011**, *1*, 826.
- (14) Limongelli, V.; Bonomi, M.; Parrinello, M. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 6358.
- (15) Wood, Z. A.; Poole, L. B.; Karplus, P. A. *Science* **2003**, *300*, 650.
- (16) Knoops, B.; Goemaere, J.; Van der Eecken, V.; Declercq, J. P. *Antioxid. Redox Signal.* **2011**, *15*, 817.
- (17) Wood, Z. A.; Schroder, E.; Robin Harris, J.; Poole, L. B. *Trends Biochem. Sci.* **2003**, *28*, 32.
- (18) Smeets, A.; Marchand, C.; Linard, D.; Knoops, B.; Declercq, J. P. *Arch. Biochem. Biophys.* **2008**, *477*, 98.
- (19) Declercq, J. P.; Evrard, C.; Clippe, A.; Stricht, D. V.; Bernard, A.; Knoops, B. *J. Mol. Biol.* **2001**, *311*, 751.
- (20) Hall, A.; Parsonage, D.; Poole, L. B.; Karplus, P. A. *J. Mol. Biol.* **2010**, *402*, 194.
- (21) Aguirre, C.; ten Brink, T.; Guichou, J. F.; Cala, O.; Krimm, I. *PLoS One* **2014**, *9*, e102300.
- (22) Laio, A.; Parrinello, M. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 12562.
- (23) Barducci, A.; Bussi, G.; Parrinello, M. *Phys. Rev. Lett.* **2008**, *100*, 020603.
- (24) Limongelli, V.; De Tito, S.; Cerofolini, L.; Fragai, M.; Pagano, B.; Trotta, R.; Cosconati, S.; Marinelli, L.; Novellino, E.; Bertini, I.; Randazzo, A.; Luchinat, C.; Parrinello, M. *Angew. Chem.* **2013**, *52*, 2269.
- (25) Limongelli, V.; Marinelli, L.; Cosconati, S.; La Motta, C.; Sartini, S.; Mugnaini, L.; Da Settimo, F.; Novellino, E.; Parrinello, M. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 1467.
- (26) Grazioso, G.; Limongelli, V.; Branduardi, D.; Novellino, E.; De Micheli, C.; Cavalli, A.; Parrinello, M. *J. Am. Chem. Soc.* **2012**, *134*, 453.
- (27) Di Leva, F. S.; Novellino, E.; Cavalli, A.; Parrinello, M.; Limongelli, V. *Nucleic Acids Res.* **2014**, *42*, 5447.
- (28) Bonomi, M.; Branduardi, D.; Bussi, G.; Camilloni, C.; Provasi, D.; Raiteri, P.; Donadio, D.; Marinelli, F.; Pietrucci, F.; Broglia, R. A.; Parrinello, M. *Comput. Phys. Commun.* **2009**, *180*, 1961.
- (29) Fielding, L. *Prog. Nucl. Magn. Reson. Spectrosc.* **2007**, *51*, 219.
- (30) Markin, C. J.; Spyropoulos, L. *J. Biomol. NMR* **2012**, *53*, 125.
- (31) Barelier, S.; Linard, D.; Pons, J.; Clippe, A.; Knoops, B.; Lancelin, J. M.; Krimm, I. *PLoS One* **2010**, *5*, e9744.
- (32) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.
- (33) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins* **2006**, *65*, 712.
- (34) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. *Proteins* **2010**, *78*, 1950.
- (35) Jorgensen, W. L.; Madura, J. D. *J. Am. Chem. Soc.* **1983**, *105*, 1407.
- (36) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157.
- (37) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. *J. Mol. Graphics Modell.* **2006**, *25*, 247.
- (38) Feenstra, K. A.; Hess, B.; Berendsen, H. J. C. *J. Comput. Chem.* **1999**, *20*, 786.
- (39) Humphrey, W.; Dalke, A.; Fausch, S.; Schulten, K.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33.
- (40) Gasteiger, E.; Hoogland, C.; Gattiker, A.; Duvaud, S. e.; Wilkins, M.; Appel, R.; Bairoch, A. In *The Proteomics Protocols Handbook*; Walker, J., Ed.; Humana Press: New York, 2005; p 571.
- (41) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Pfeifer, J.; Bax, A. *J. Biomol. NMR* **1995**, *6*, 277.
- (42) Johnson, B. A.; Blevins, R. A. *J. Biomol. NMR* **1994**, *4*, 603.
- (43) Farmer, B. T.; Constantine, K. L.; Goldfarb, V.; Friedrichs, M. S.; Wittekind, M.; Yanchunas, J.; Robertson, J. G.; Mueller, L. *Nat. Struct. Mol. Biol.* **1996**, *3*, 995.
- (44) Schumann, F.; Riepl, H.; Maurer, T.; Gronwald, W.; Neidig, K.-P.; Kalbitzer, H. *J. Biomol. NMR* **2007**, *39*, 275.
- (45) Lange, A.; Ismail, M. B.; Riviere, G.; Hologne, M.; Lacabanne, D.; Guilliere, F.; Lancelin, J. M.; Krimm, I.; Walker, O. *FEBS Lett.* **2012**, *586*, 3379.

Funnel-Metadynamics and Solution NMR to Estimate Protein-Ligand Affinities

Laura Troussicot[§], Florence Guillière[§], Vittorio Limongelli^{¶#}, Olivier Walker[§] and Jean-Marc Lancelin^{§*}.

[§] Université de Lyon, Institut des Sciences Analytiques, UMR 5280, CNRS, Université Lyon 1, ENS Lyon - 5, rue de la Doua, F-69100 Villeurbanne, France. [¶] Università della Svizzera Italiana (USI), Faculty of Informatics, Institute of Computational Science, via G. Buffi 13, CH-6900 Lugano, Switzerland. [#] Department of Pharmacy, University of Naples "Federico II", via D. Montesano 49, I-80131 Naples, Italy.
KEYWORDS: Funnel metadynamics, solution NMR, absolute binding free energy, Human peroxiredoxin 5.

SUPPORTING INFORMATION

Figure S1

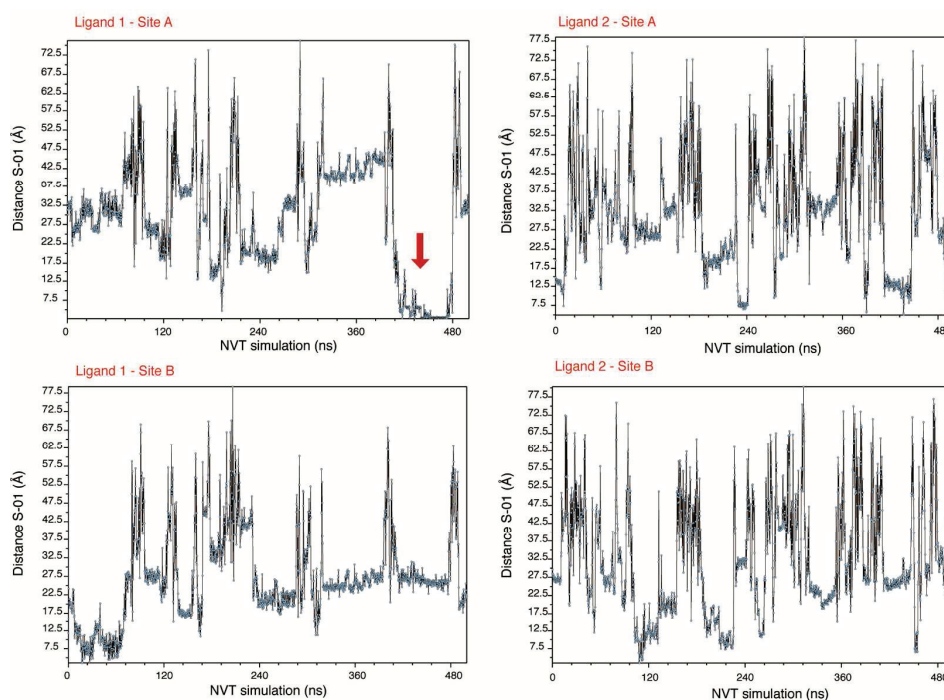


Figure S1 (continued)

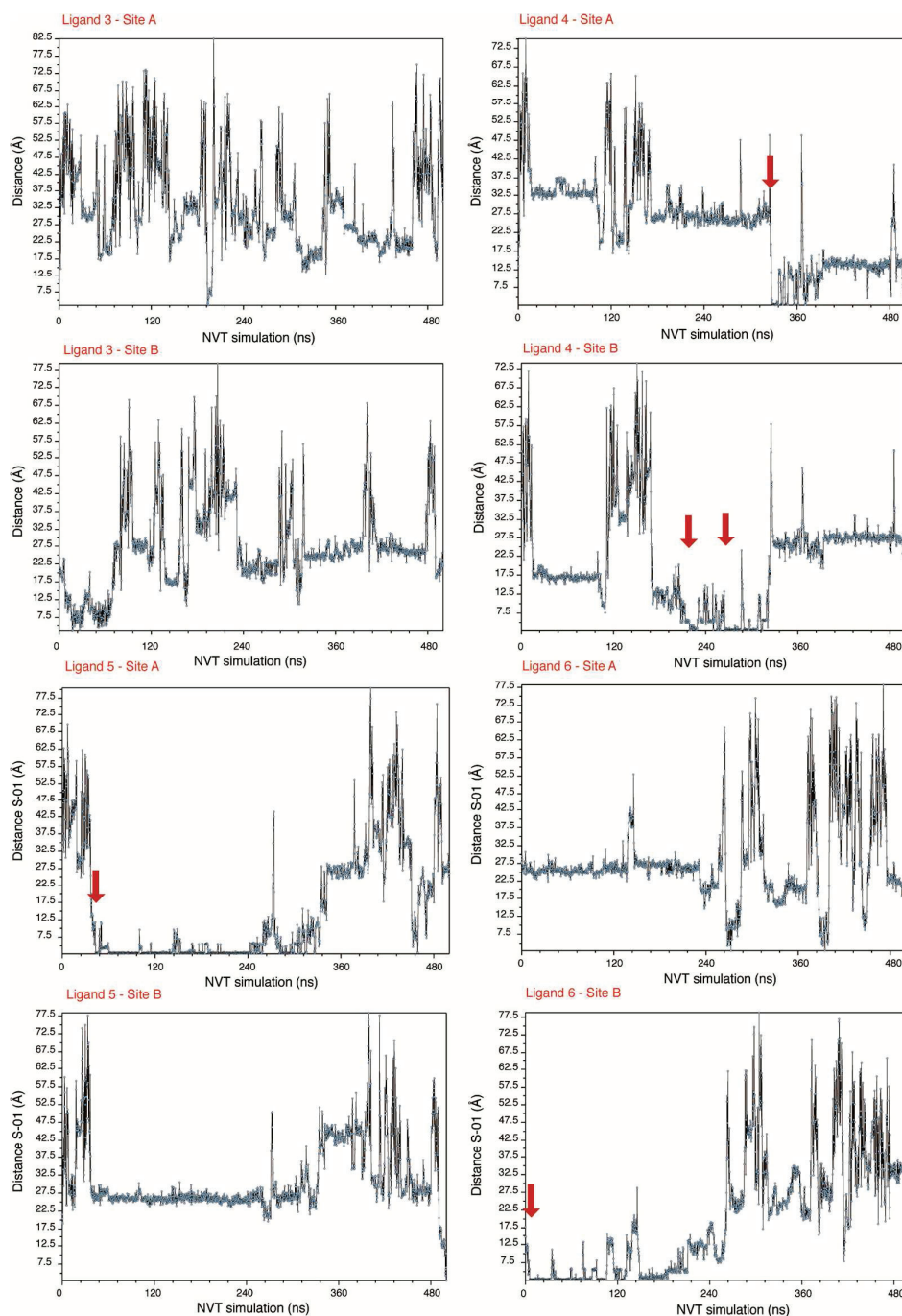


Figure S1. NVT ligand trajectories during an unbiased NVT 500 ns dynamics into explicit solvent at high-concentration of 4-methylcatechol in the presence of equilibrated Human peroxiredoxin PRX5 crystal structure (PDB entry 3MNG.pdb²⁰) in the TIP3P water model for the explicit solvent model³⁵. Protein-ligand distances between the thiolate sulfur atom of each active sites of the homodimer (site A and site B) and the oxygen atom O1 of 4-methylcatechol are plotted against the NVT simulation time for each of the six 4-methylcatechol molecules (ligands 1 to 6). Red arrows indicate the beginning of a ligand-binding period to the active site of PRX5.

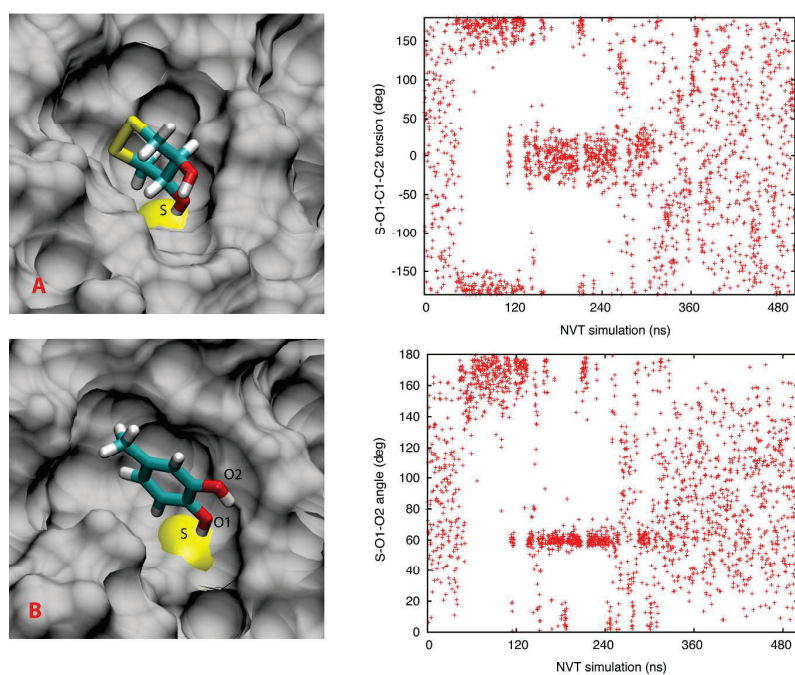


Figure S2. Ligands bound conformations to the active site of Human peroxiredoxin PRX5 and NVT trajectory during an unbiased NVT 500 ns dynamics into explicit solvent at high-concentration of 4-methylcatechol. *A*: snapshot at 2ns of NVT of oxidized dithiothreitol (DTTox) H-bonded in the active site starting from the equilibrated crystal structure (PDB entry 3MNG.pdb²⁰) in the TIP3P water model for the explicit solvent model³⁵. The surface accessible to the solvent of PRX5 is represented in gray and the van der Waals sphere of Cys47 sulfur atom thiolate is drawn in yellow. *B*: Snapshot of 4-methylcatechol H-bonded in the active site at 90 ns of NVT (ligand 5 in Fig. S1). Protein to 4-methylcatechol ligand Cys47 S...O1-C1-C2 torsion (upper-right panel) and Cys47 S...O1-O2 angle is plotted along the NVT simulation time every 240 ps of dynamics. 4-Methylcatechol (ligand 5 of Fig. S1) binds first in the conformation shown in *B*, then exchanged with 2 H-bond conformers before unbinding (see the text).

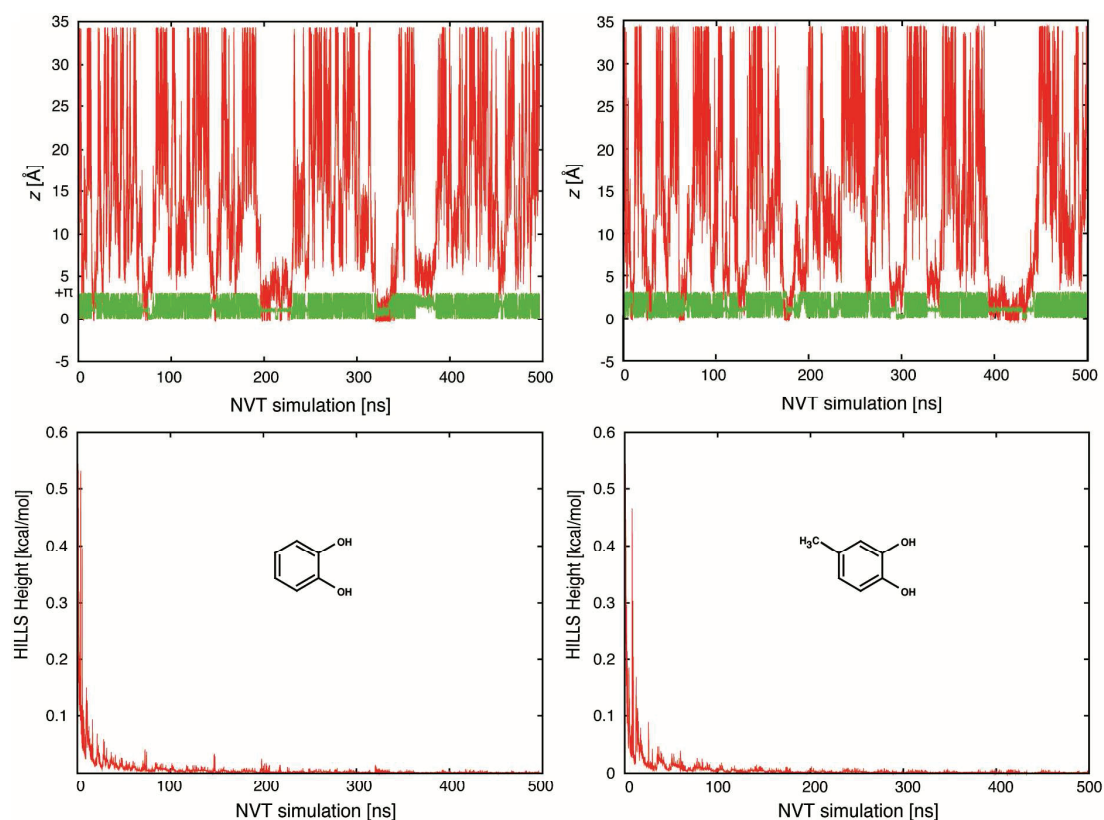


Figure S3. Funnel metadynamics trajectories of Human peroxiredoxin PRX5 in the presence of catechol (left panels) of 4-methylcatechol (right panels). Upper panels, red traces : position of the ligand projected on the funnel Z-axis in Å (CV₁) as a function of the simulation time. The ligands sample the whole Z-axis dimension with the defined 0-34 Å limits of the fixed funnel external potential. The green traces indicate the collective variable Cys₄₇ S...O₁-O₂ angle (CV₃) reported in radian. Above Z-axis=18 Å, the ligand samples all conformations in the cylinder section of the funnel and interacts only with the solvent. Below 18 Å, the ligand interacts with the protein in the region of the active site. Binding periods correspond to the defined CV₃ and the short CV₁. Lower panels, HILLS potentials deposition in kcal/mol along the simulation time indicating the progressing convergence where no energy depositions occur anymore.

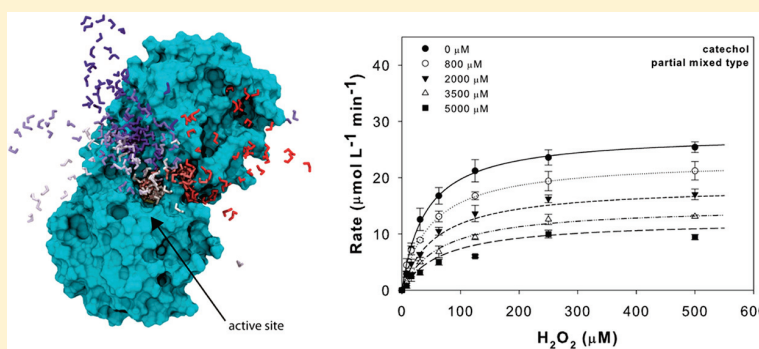
Predicting and Understanding the Enzymatic Inhibition of Human Peroxiredoxin 5 by 4-Substituted Pyrocatechols by Combining Funnel Metadynamics, Solution NMR, and Steady-State Kinetics

Melissa L. Chow,[†] Laura Troussicot,[†] Marie Martin,[†] Bastien Doumèche,[‡] Florence Guillière,[†] and Jean-Marc Lancelin^{*,†}

[†]Université de Lyon, Institut des Sciences Analytiques, UMR 5280, CNRS, Université Lyon 1, ENS Lyon-5, rue de la Doua, F-69100 Villeurbanne, France

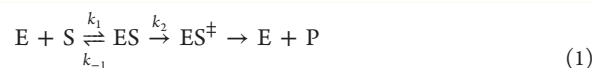
[‡]Université de Lyon, Institut de Chimie et Biochimie Moléculaires et Supramoléculaires, UMR 5246, Université Claude Bernard Lyon 1, CNRS, INSA, CPE-Lyon, 43 Bd du 11 novembre 1918, F-69622 Villeurbanne, France

S Supporting Information



ABSTRACT: Funnel metadynamics is a kind of computational simulation used to enhance the sampling of protein–ligand binding events in solution. By characterization of the binding interaction events, an estimated absolute binding free energy can be calculated. Nuclear magnetic resonance and funnel metadynamics were used to evaluate the binding of pyrocatechol derivatives (catechol, 4-methylcatechol, and 4-*tert*-butylcatechol) to human peroxiredoxin 5. Human peroxiredoxins are peroxidases involved in cellular peroxide homeostasis. Recently, overexpressed or suppressed peroxiredoxin levels have been linked to various diseases. Here, the catechol derivatives were found to be inhibitors against human peroxiredoxin 5 through a partial mixed type noncompetitive mechanism. Funnel metadynamics provided a microscopic model for interpreting the inhibition mechanism. Correlations were observed between the inhibition constants and the absolute binding free energy. Overall, this study showcases the fact that funnel metadynamics simulations can be employed as a preliminary approach to gain an in-depth understanding of potential enzyme inhibitors.

Enzyme catalysis is one of the more fascinating properties of biological macromolecules. Since the seminal reports of Henri, Michaelis, and Menten, it has been known that the rate of enzyme catalysis of a chemical reaction is directly proportional to the concentration of the enzyme–substrate complex. This concept was predicted by the original Henri–Michaelis–Menten equations.^{1–4} A century of enzymology and enzyme kinetics established the different classical models of the enzyme catalytic mechanisms and models for their inhibitions.⁵ In basic enzyme catalysis, an enzyme binds reversibly to its substrate and forms an enzyme–substrate complex that is slightly stabilized. From this state, the reaction evolves to a chemical transition state with a decrease in free energy and the production of a product. This reaction is represented here (eq 1):



where E is the enzyme, S its substrate, ES the enzyme–substrate complex or Michaelis–Menten complex, ES[‡] the enzyme transition-state complex, and P the product of the reaction. The complete catalytic cycle can be complicated by modifications to the enzyme after the reaction and its regeneration involving another enzyme or cofactor (coenzyme). Many types of enzymatic inhibition models are known, including sophisticated allosteric inhibitions.⁶ The simplest inhibition model is based on the affinity of an inhibitor (I),

Received: April 20, 2016

Revised: May 25, 2016

Published: May 30, 2016

analogous to the substrate, for the enzyme's active site leading to a nonproductive enzyme–inhibitor complex (EI) (eq 2).⁴



Competition between a substrate and an inhibitor for the active site is either reversible or irreversible when the inhibitor reacts by forming covalent bonds to the enzyme. In the case of reversible competitive inhibition, the greater the affinity (and the higher the stability of the enzyme–inhibitor complex), the more potent the inhibition.

Protein–ligand binding interaction is the fundamental basis for understanding any type of inhibition mechanism. Also, it is the core and the key point for designing inhibitors for developing selective drugs or effectors. Recently, the process of binding and unbinding of a ligand to a protein could be rationalized more easily using accelerated molecular dynamics protocols along with high-performance computing (HPC) at the laboratory scale.⁷ This has led to advancements in obtaining statistical descriptions of protein–ligand interactions at the microscopic level for macroscopic measurements.

Previously, Troussicot et al. demonstrated that nuclear magnetic resonance (NMR) assays of protein–ligand affinity measurement of the chemical shift perturbations were in agreement with binding affinities calculated by funnel metadynamics (FM) simulations.^{8,9} Human peroxiredoxin 5 (hPrx5), a 2×19 kDa homodimer, was determined to weakly but selectively bind to catechol and 4-methylcatechol within the active site.⁸ hPrx5 is part of the peroxiredoxin (Prx) family that reduces and detoxifies hydrogen peroxide, peroxynitrites, and a range of hydroperoxides (ROOH). Prxs are antioxidant defense enzymes along with superoxide dismutase (SOD), catalase, and glutathione peroxidase (GPX).^{10,11} Antioxidant enzymes provide protection from reactive oxygen species (ROS) such as superoxide radical anions, hydrogen peroxide (H_2O_2), and hydroxyl radicals.¹⁰ Additionally, Prxs are involved in regulating the levels of peroxides in cells linked to both stress- and non-stress-related signaling.^{11,12}

There are six mammalian Prx members that can be organized into three classes: typical 2-Cys (Prx1–4), atypical 2-Cys (Prx5), and 1-Cys (Prx6) reflecting the number of cysteine residues involved in the redox reaction.^{13,14} Prx catalytic activity reduces H_2O_2 to water. The first step is common for all classes: the active site is fully folded (FF) with a peroxidatic cysteine (Cp) that attacks and reacts with hydrogen peroxide. Cp (Cys-S⁻) is then oxidized to sulfenic acid (Cys-SOH) and produces water or alcohols.^{13,15} The active site of Prx then becomes locally unfolded (LU), and Prx either can be overoxidized to sulfinic acid (Cys-SO₂H) after reacting with another H_2O_2 becoming inactivated or can be regenerated back to Cys-SH. The recycling of reduced Prx depends on the Prx class.^{12,15} The most common classes are 2-Cys (typical and atypical), in which the oxidized Cys-SOH reacts with the resolving cysteine (Cr) from the same subunit or another subunit forming a disulfide bond. Cp–Cr bond formation leads to the rearrangement of the active site. Lastly, the disulfide bond is reduced by thioredoxin (Trx), by a thioredoxin-like protein, or by glutathione (GSH), returning Prx to the Cys-SH state with an FF active site.^{10,11,16}

Oxidative stresses from the ROS have been linked to the development of many diseases and aging, which result in the impairment of several cellular functions.¹⁶ Prxs have been linked to cancers by influencing signaling cascades, growth control, tumor suppression, and chemotherapy resistance.¹⁷

Also, Prxs were identified in playing a role during the inflammation cascade in the brain after an ischemic stroke.¹⁸ Deficiencies of Prxs in prokaryotic and eukaryotic pathogens have been linked to compromising host immune systems leading to infections. Prxs have been associated with various diseases, and therefore, there is an interest in identifying drugs that target Prxs for therapeutic treatments.^{11,15} There have been some advancements in identifying drugs for inhibiting or modulating Prxs activity such as conoidin A,¹⁹ adenanthin,²⁰ H7,²¹ triptolide, celastrol, withaferin A,²² and disulfide BNP7787.²³

Here, NMR and FM simulation experiments were used to characterize the binding and inhibition of hPrx5 by the three pyrocatechol (or catechol) derivatives: catechol, 4-methylcatechol, and 4-*tert*-butylcatechol. FM was used to provide details at the microscopic level, specifically for the binding interaction between hPrx5 with hydrogen peroxide and the three ligands within the active site. Additionally, the binding interaction between a mimicked ES complex and the ligands was also assessed. Evaluating these binding states contributed to there being more information about the binding preference among the catechol derivatives binding for hPrx5. Also, solution NMR experiments were able to determine that hPrx5 has binding affinity for 4-*tert*-butylcatechol, and the ligand bound to the hPrx5 active site in a manner similar to that of catechol and 4-methylcatechol as previously determined. Furthermore, the potential for the catechol derivatives to bind and inhibit hPrx5 catalytic activity was evaluated. The *in vitro* enzyme assays proved that all the catechol derivatives can inhibit hPrx5 catalytic activity and inhibit via a partial mixed type non-competitive mechanism. Overall, this study highlights the prospects of FM being utilized as a method for predicting ligands as inhibitors assisting in drug design.

■ MATERIALS AND METHODS

System Preparation and Equilibration for MD. Systems subjected to MD protocols were prepared as previously described⁸ starting from the conformation of the high-resolution crystal structure of homodimeric hPrx5 corresponding to Protein Data Bank (PDB) entry 3MNG.^{24,25} Only the crystallographic water molecules were retained from the PDB coordinate file in addition to the protein, and the rest were eliminated. Hydrogen peroxide and/or ligands were positioned manually near one of the two active sites of the homodimer. Simulations were conducted using the AMBER99SB-ILDN force field^{26–28} for the protein and the TIP3P water model for the explicit solvent.²⁹ Active site Cys47 was modeled as a thiolate (residue CYM of Amber). Na⁺ and Cl⁻ were added to match 150 mM sodium chloride aqueous solutions, close to the experimental conditions used. The Amber charges were applied to the protein, substrate, ligands, ions, water atoms, and the restrained electrostatic potential charges were used for the ligands using the Antechamber program suite³⁰ and the General Amber Force Field GAFF.³¹ Using ACEMD code,³² the system was minimized and equilibrated under constant-pressure and -temperature (NPT) conditions at 1 atm and 300 K using a time step of 4 fs because of the use of the hydrogen mass repartitioning scheme³³ implemented in ACEMD, with a nonbonded cutoff of 9 Å, rigid bonds, and particle mesh Ewald long-range electrostatics with a grid with spacing of 1 Å. The systems were equilibrated first using 500 steps of steepest-descent minimization, followed by running the isothermal NVT ensemble for 0.1 ns, using a Langevin thermostat set at 300 K,

followed by 5 ns of the isothermal–isobaric *NPT* ensemble using a Langevin thermostat at the same temperature and the Berendsen barostat of ACEMD. During minimization and equilibration, the heavy protein and ligands atoms were restrained spatially using a spring constant of 10 kcal mol⁻¹ Å⁻². The magnitude of the restraining force constant was then reduced to 1 kcal mol⁻¹ Å⁻² during 100 ps of *NVT*. The barostat was switched on at 1 atm for 1 ns of *NPT* simulation. During that period, the force constant of the position constraint of all heavy atoms was gradually reduced every 100 ps steps by a factor of 0.65 to a final value of ~0.1 kcal mol⁻¹ Å⁻², allowing the systems to relax smoothly. Finally, the volume was allowed to relax for a further 4 ns under *NPT* conditions, reaching a final box size. During this run, *Cα* atoms of the protein and ligands were restrained with a 1 kcal mol⁻¹ Å⁻² harmonic potential to prevent the system from reorienting. Then, the equilibrated systems were used for MD.

Funnel Metadynamics. The PLUMED plugin³⁴ was used to conduct funnel metadynamics calculations. Because the funnel external potential is fixed in space,⁹ the hPrx5 dimer is diffusion-restrained by three atoms chosen far from the dimer interface and the active site as previously described.⁸ *Cα* atoms of G6, G31, and K65 of chain A were used as positional restrains to inhibit the whole protein diffusion with a force constant of 20 kcal mol⁻¹ Å⁻² in ACEMD. The bias was added on a distance CV (see Results). A Gaussian width of 0.35 Å was used, and a Gaussian deposition rate of 0.25 kcal mol⁻¹ ps⁻¹ (1 kcal = 4.18 kJ) was initially used and gradually decreased on the basis of the adaptive bias with a ΔT of 3600 K. Trajectories and geometrical clustering were analyzed using the VMD software.³⁵ In the FM calculation with hydrogen peroxide bound to hPrx5, a wall potential was used with PLUMED from the sulfur to the hydrogen of hydrogen peroxide of 20 kcal mol⁻¹ Å⁻² applying beyond a maximal interatomic distance of 2.0 Å. This term is not taken into account in the free energy calculations from PLUMED outputs.

Protein Expression and Purification. Recombinant human peroxiredoxin 5 (hPrx5, EC 1.11.1.15), yeast thioredoxin 1 (yTrx1, EC 1.8.1.9), and yeast thioredoxin reductase 1 (yTrxR1, EC 1.8.1.9) were in the pQE-30 expression vector (Qiagen) with an N-terminal fusion with a hexahistidine (6xHis) tag. The plasmids were provided through collaboration with B. Knoop's lab at the Université Catholique de Louvain (Louvain-la-Neuve, Belgium). The expression and purification of the proteins were adapted from ref 36. Briefly, the plasmids were transformed into *Escherichia coli* strain M15 (pRep4) and grown in LB medium at 37 °C with 50 µg/mL ampicillin and kanamycin. For the production of [¹⁵N]hPrx5, M9 minimal medium was used (6 g/L Na₂HPO₄, 3 g/L KH₂PO₄, and 0.5 g/L NaCl) supplemented with 50 µg/mL ampicillin and kanamycin, 6 µg/L thiamine, 1 mM MgSO₄, 1 mM CaCl₂, a 10 mL/L trace metal solution [5 g of EDTA, 0.5 g of FeCl₃, 0.005 g of ZnO, 0.001 g of CuCl₂, 0.001 g of Co(NO₃)₂·6H₂O, and 0.001 g of (NH₄)₆Mo₇O₂₄·4H₂O per liter], 4 g/L D-glucose, and 1 g/L ¹⁵NH₄Cl.³⁷ The bacterial cultures were induced at an O.D.600 of 0.6 with 1 mM isopropyl thiogalactopyranoside (IPTG) for 4–5 h. The bacterial cells were harvested by centrifugation for 20 min at 2000g [Thermo Fisher Scientific, F15-8 × 50 cy rotor]. The bacterial cell pellets were lysed with 10 mM imidazole, 50 mM phosphate, and 300 mM NaCl (pH 8) by sonication (Sartorius Stedim Biotech, Aubagne, France). The cell lysate was then centrifuged for 45 min at 16500g (Thermo Fisher Scientific, TX-750 rotor). The

cell supernatant was then loaded onto a Ni²⁺-NTA column (Qiagen). The column was washed with the lysis buffer. The protein was eluted and dialyzed overnight against PBS (pH 7.4) at 4 °C. The homogeneity of the protein was analyzed by sodium dodecyl sulfate–polyacrylamide gel electrophoresis and the purified hPrx5 stored at –80 °C until it was used. The protein concentration was quantified by UV spectrometry (UV–vis spectrometer, Jasco France) using the following protein extinction coefficients at A₂₈₀ (ϵ calculated via expasy.org/protparam): for yTrx1, $\epsilon = 10095 \text{ M}^{-1} \text{ cm}^{-1}$; for yTrxR1, $\epsilon = 24660 \text{ M}^{-1} \text{ cm}^{-1}$; and for hPrx5, $\epsilon = 5625 \text{ M}^{-1} \text{ cm}^{-1}$.

NMR Assays. The binding of hPrx5 to the catechol derivatives was examined through two-dimensional (¹H–¹⁵N) NMR experiments, also termed heteronuclear single-quantum coherence (HSQC). The experimental procedure was adapted from Troussicot et al.⁸ with hPrx5 amino acid residues previously assigned. Briefly, 200 µM ¹⁵N-labeled hPrx5 reduced (2 mM TCEP) in PBS (pH 7.4) was exposed to increasing concentrations of catechol derivatives (in 110 mM DMSO or H₂O) until binding saturation was exhibited. The binding interactions were monitored at 28 °C on a Varian Inova 600 MHz NMR spectrometer equipped with a 5 mm standard triple-resonance (¹H/¹³C/¹⁵N) inverse probe with a z-axis field gradient. The ¹H–¹⁵N spectra were processed using the NMRPipe/NMDraw package.³⁸ The chemical shift perturbations (CSPs) were assigned using NMRView.³⁹ The combined amide CSP $\Delta\delta$ was defined according to the following expression⁴⁰ (eq 3):

$$\Delta\delta = \sqrt{\frac{\Delta\delta_{\text{H}}^2 + \frac{1}{25}\Delta\delta_{\text{N}}^2}{2}} \quad (3)$$

Upon characterization of the amide CSP $\Delta\delta$ for the specific amino acid residues involved in the binding interactions, a binding dissociation constant (K_{D}) was extrapolated from the plot of CSPs versus [catechol derivative]/[hPrx5] as previously described⁸ using MatLab (Mathworks, Meudon, France). Data were acquired in duplicate.

Half-Maximal Inhibitory Concentration Assay (IC₅₀). The inhibitory concentration assay of hPrx5 was adapted from Kim et al.⁴¹ and Barelier et al.³⁷ The inhibition activity of hPrx5 was measured indirectly by the coupling reaction of yTrxR1 oxidizing NADPH, monitored by the absorbance at 340 nm as shown in Figure 7. The reaction conditions consisted of catechol derivatives at various concentrations (catechol at 0–75 mM, 4-methylcatechol at 0–15 mM, and 4-*tert*-butylcatechol at 0–5 mM) in PBS, 15 µM yTrx, 2 µM yTrxR, 0.15 µM hPrx5, 200 µM NADPH (Sigma-Aldrich), and 25 µM H₂O₂ (at approximately the K_{M} value) in a total reaction volume of 1000 µL. The reaction was conducted with a mixture of all three proteins diluted in PBS [10 mM NaHPO₄, 137 mM NaCl, and 3 mM KCl] with 1 mM EDTA at pH 7.0 upon addition of the catechol derivatives and NADPH. Lastly, H₂O₂ was added to the mixture and the solution mixed to initiate and monitor the inhibition of hPrx5 at room temperature (23 °C). The reaction was monitored at A₃₄₀ for 200 s, and the initial rate of the reaction was determined from the linear portion of the curve and was expressed in micromoles per liter per minute in reference to the amount of hPrx5 used. The IC₅₀ values were determined by plotting the reaction rates versus the catechol derivative concentration and fitting the data to a hyperbolic decay equation using SigmaPlot. The catechol derivatives

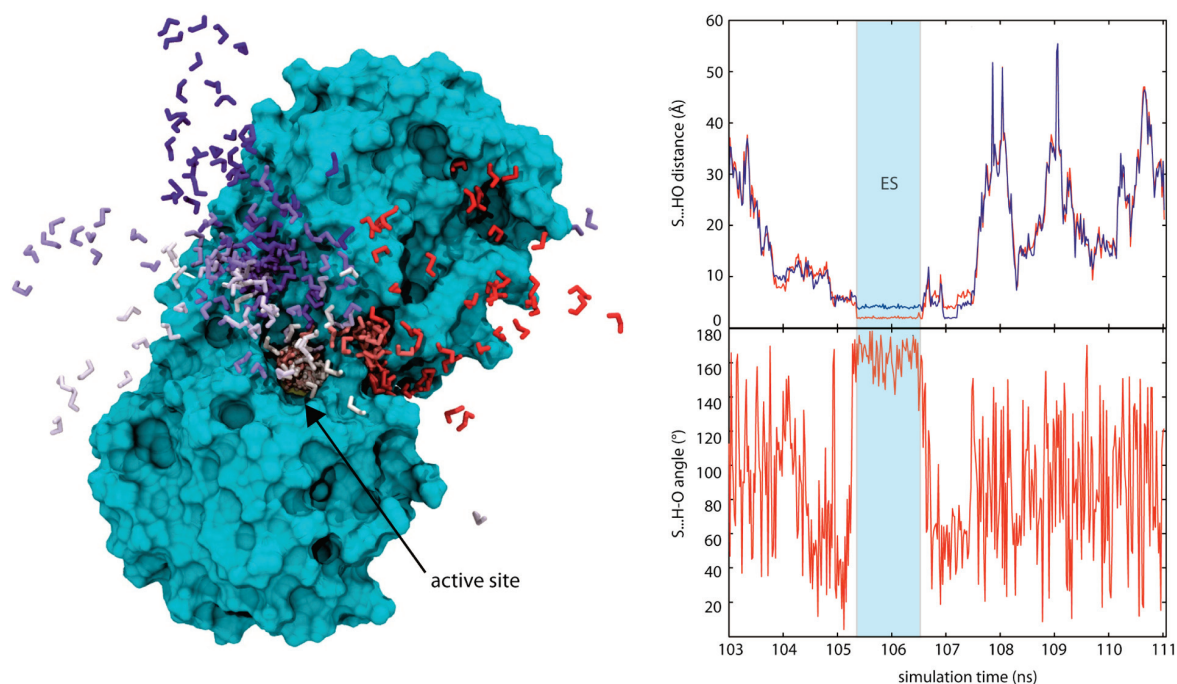


Figure 1. Illustration of a binding event between hydrogen peroxide and hPrx5 simulated by unrestrained NVT molecular dynamics over 1.2 ns between 105 and 107 ns of a 350 ns unrestrained NVT trajectory. In the left panel, hPrx5 is represented with the water-accessible surface colored light blue and hydrogen peroxide with bonds colored using a deep blue–white–red continuous spectrum depending of the simulation time. The time step between two frame shots is 20 ps. The hydrogen peroxide bound to hPrx5 corresponds to the Michaelis–Menten complex colored white in one of the two active sites of the homodimer (indicated with an arrow). The right panels show interatomic distances (top) from the Cys47-S[−] thiolate to a hydrogen of H₂O₂ (blue and red traces) as a function of simulation time. The bond angle is defined by the sulfur atom of the Cys47 thiolate, the one hydrogen atom of the hydrogen peroxide, and the oxygen atom attached (plotted in the bottom panel). The binding event is depicted in the blue highlighted region.

inhibitory potency against hPrx5 was ranked based on their IC₅₀ values. The data are representative of three individual experimental days ($n = 3$).

Inhibition Mechanism Assay. The inhibition mechanism assay was conducted with three or four concentrations of the catechol derivatives above and below the IC₅₀ value, 15 μ M γ Trx, 2 μ M γ TrxR, 0.15 μ M hPrx5, 200 μ M NADPH, and a range of H₂O₂ concentrations from 0 to 500 μ M. The reaction was initiated by mixing all three proteins together with a catechol derivative and NADPH and adding H₂O₂. The reaction was monitored by measuring the A₃₄₀ for 200 s, and the initial rate of the reaction was determined from the linear portion of the curve and is expressed in micromoles per milligram per minute. Data were plotted in a Michaelis–Menten and/or Hanes–Woolf representation.⁴ The V_{\max} and K_M values were calculated for each inhibitor concentration using SigmaPlot. Since the data sets were not ideally linear, DynaFit software was used with the steady-state approximation for model discriminations and parameter determination.^{42,43} The data are triplicate results from three individual experiments ($n = 3$).

RESULTS

Unbiased NVT Molecular Dynamics Simulating Binding of H₂O₂ to Reduced hPrx5. To study how hydrogen peroxide (H₂O₂) binds to reduced human peroxidoredoxin 5 (hPrx5) and to deduce the pertinent collective variables for metadynamics simulations, the reduced hPrx5 system was first equilibrated in the presence of 6 mM H₂O₂ (i.e., six molecules

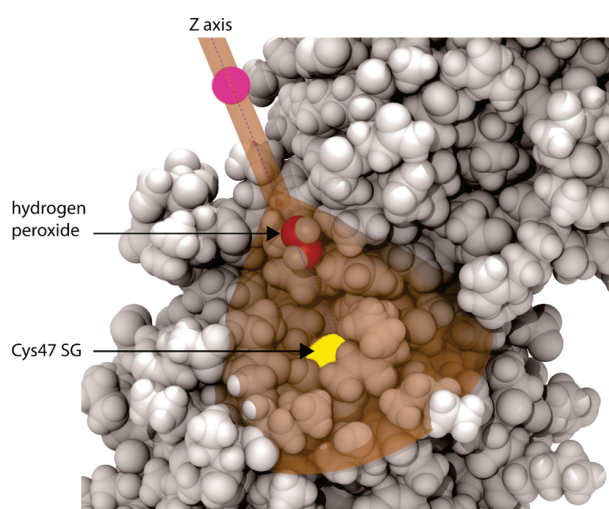


Figure 2. External funnel restraint potential (transparent orange surface) applied to the hPrx5 active site with its substrate, hydrogen peroxide. The catalytic thiolate of Cys47 (SG) is represented with a yellow sphere. The cylinder part of the funnel is aligned and centered on the Z-axis with a radius R_{cyl} of 1 Å. The funnel potential is limited to 35 Å from the Z-axis origin (SG atom of Cys47) by a wall potential represented by a pink sphere on the top. The cone region is defined by a vertex height Z_{cc} of 18 Å from the origin and an angle α of 1.1 rad.

of H₂O₂ were introduced around the protein in the solvation cell that includes \sim 55000 water molecules) before a run of 350

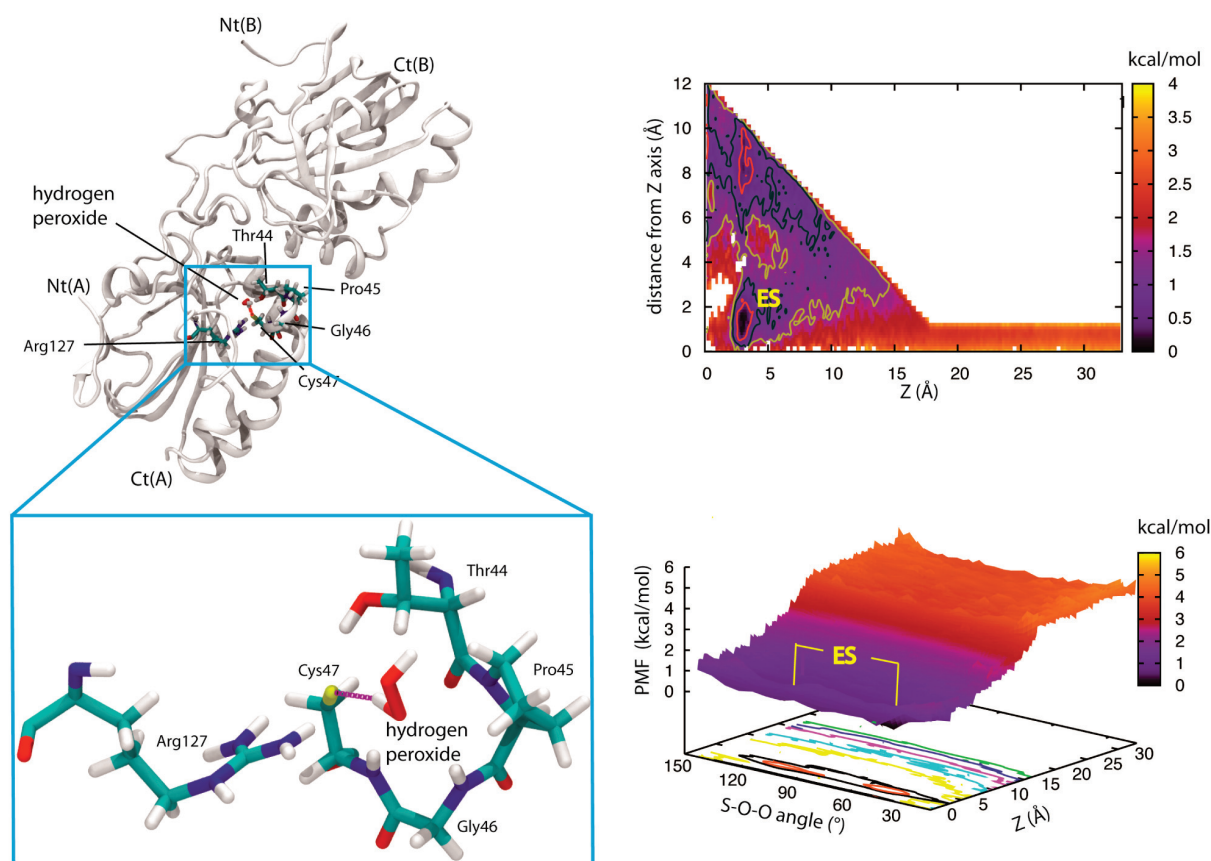


Figure 3. Michaelis–Menten complex (ES) of hPrx5 and hydrogen peroxide determined using funnel metadynamics (FM). The top left panel is a structure of hPrx5 complexed with hydrogen peroxide during FM. The active site is located at the interface between chains A and B of the homodimer. The main amino acids participating in hydrogen peroxide reduction (Thr44, Pro45, Gly46, Cys47 thiolate, and Arg127) are represented as cyan sticks. N- and C-termini of each chain are denoted as Nt (A or B) and Ct (A or B), respectively. The bottom left panel is an enlarged view of ES corresponding to hydrogen peroxide H-bonded to the thiolate of Cys47. The H-bond is indicated with a purple dotted line. H-Bond formations fluctuated along the MD trajectory, alternating between Arg127 or Thr44 as the donor and Cys47 thiolate as the acceptor. The right panels show free energy surfaces (FES) represented as a function of the three collective variables used in FM calculations. The potential mean force (PMF) is mapped with iso-energy contours using $0.5 \text{ kcal mol}^{-1}$ intervals.

ns of NVT molecular dynamics (MD). H_2O_2 molecules were initially dispersed at arbitrary positions around the protein surface from 20 to 25 Å. The NVT MD trajectory was analyzed for possible binding events between H_2O_2 and hPrx5, by searching for hydrogen bonds (H-bonds) formed between the active site amino acids, in particular involving Cys47 thiolate and H_2O_2 . Three short binding events were detected for H_2O_2 bound within the two active sites of hPrx5 as a homodimer. The longest binding event observed was ~ 1.5 ns. During a binding event, H_2O_2 made a H-bond with the sulfur atom of Cys47 with a H–S distance of 1.8–2.0 Å. The S \cdots H–O bond angle is $\sim 180^\circ$ (Figure 1).

Hydrogen Peroxide Funnel Metadynamics. To enhance the sampling of infrequent events (i.e., occurring on a long time scale) in a reasonable computation time, well-tempered metadynamics was proposed. Metadynamics provides a free energy landscape of the investigated processes.^{44,45} Free energy maps can be computed at the end of the calculation using the history-dependent bias potential added during the simulation with a few degrees of freedom of the system, called collective variables (CVs). FM provided a quantitative estimate of the binding event yielding an absolute binding free energy

difference (ΔG_b°) value between the free solvated state and the bound forms of any ligands. Using FM, the binding constant, K_b , can be computed as eq 4:

$$K_b = \pi R_{\text{cyl}}^2 \int_{\text{site}} dz \times e^{-\beta[W(z) - W_{\text{ref}}]} \quad (4)$$

where πR_{cyl}^2 is the surface of the cylinder used as the external restraint fixed potential. Potential $W(z)$ and its value in the unbound state, W_{ref} are derived from the potential mean force (PMF) obtained through metadynamics calculations. β is a constant, where $\beta = (k_B T)^{-1}$ (k_B is Boltzmann's constant and T the temperature of the system). Equilibrium binding constant K_b is directly proportional to the absolute protein–ligand binding free energy, ΔG_b° , following eq 5:

$$\Delta G_b^\circ = \frac{1}{\beta} \ln(C^\circ K_b) \quad (5)$$

where $C^\circ = 1/1660 \text{ \AA}^{-3}$ represents the standard concentration of 1 M. The equilibrium binding constant, K_b , is the inverse of the dissociation equilibrium constant, K_D , of the ligand–protein complex measured here during NMR experiments (see below).

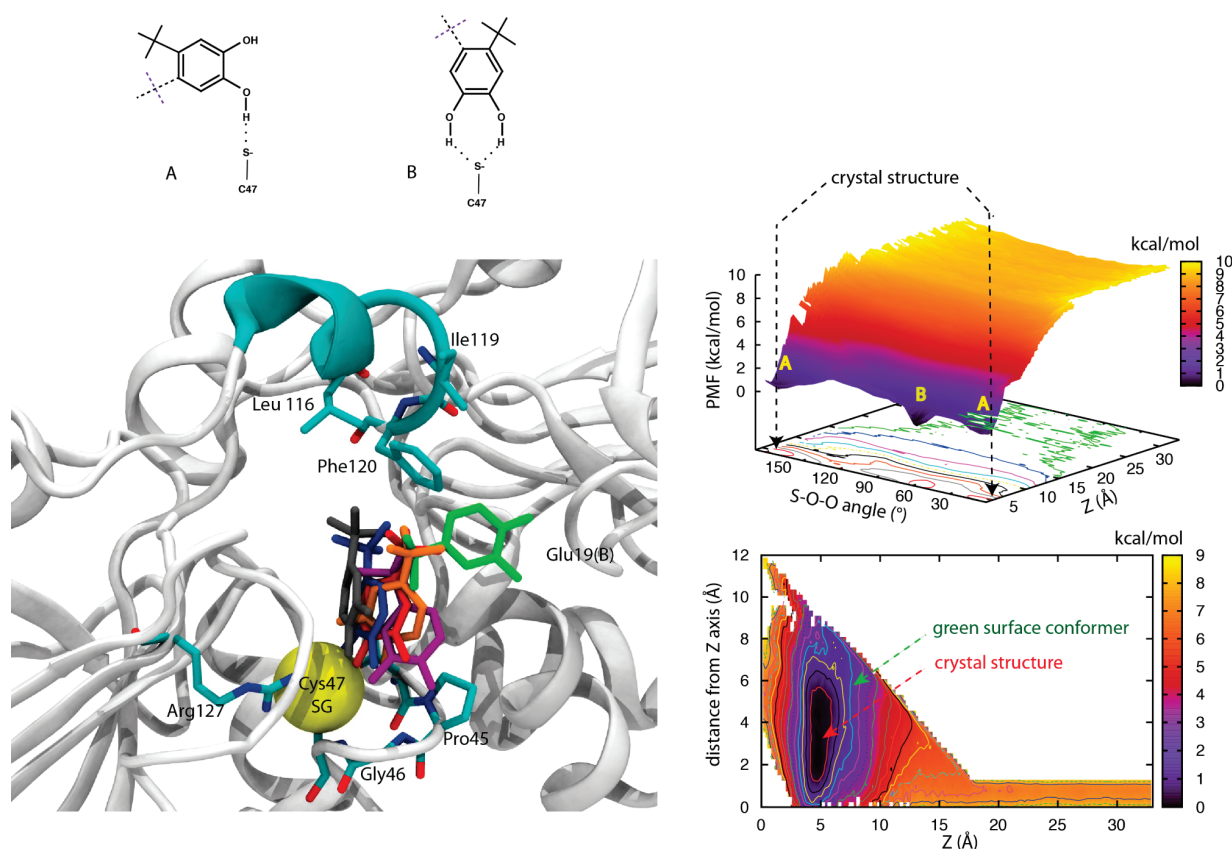


Figure 4. Binding a 4-*tert*-butylcatechol to the hPrx5 active site determined using FM. The top left panel is a schematic of the bound conformations for the ligand to Cys47 thiolate involving H-bonds. Alternative conformations are indicated with dashed bonds for the *tert*-butyl group. The bottom left panel shows six representative bound conformations (dark gray, navy blue, red, orange, purple, and green) after geometrical clustering of the whole FM trajectory. The Cys47 SG atom is represented as a yellow van der Waals sphere. Invariant amino acids Pro45, Gly46, and Arg127 are represented as cyan sticks as well as Leu116, Ile119, and Phe120 located in a two-turn helix (cyan) capping the active site. The right panels show the FES represented as a function of the three collective variables used in FM calculations. Conformations of type A and B and the crystal structure are located on the top FES. Contours of iso-energy are drawn on the map using a 1 kcal mol⁻¹ interval (top FES) or a 0.5 kcal mol⁻¹ interval (bottom FES) with the positions of the conformer indicated in green and the crystal structure indicated in red.

The funnel was aligned to the Z-axis, which was oriented to the solvent with respect to the direct access to the active site of hPrx5 (see Figure 2). The FM simulation started with H₂O₂ bound to hPrx5, and afterward, 11 successive rebinding events occurred for H₂O₂ during 200 ns of FM simulation (Figure S1). The binding of the hPrx5 substrate, hydrogen peroxide, converged to a nearly zero free energy difference and indicated that H₂O₂ diffuses freely from the solvent to the active site with a negligible affinity (Figure S2). Figure 3 (top right) shows only a tiny basin of low energy attributed to the interaction within the active site standing for ES in the simulation. The energy basin cannot be easily distinguished from other nonspecific surface interactions.

When the binding events were geometrically clustered, a single H-bond was detected between the H₂O₂ as a donor and the Cys47 thiolate. The apparent ES complex in the FM simulation is equivalent to the bound form observed in the unbiased NVT trajectories and is characterized by a geometric arrangement between hPrx5 and H₂O₂ shown in Figure 3 (bottom right). The free energy surfaces (FES) plot shows the free energy basin between H₂O₂ and hPrx5, and the bond angle conformation of the two oxygen atoms of H₂O₂ with the Cys47 sulfur exhibited an ~45° bond angle (and ~110° due to H₂O₂ symmetry).

Binding Affinity of 4-*tert*-Butylcatechol for the Free Reduced Form of hPrx5 Determined Using FM. The converged FM trajectory (Figures S3 and S4) showed that 4-*tert*-butylcatechol bound to hPrx5 with an enhanced ΔG_b° of at least -1 kcal mol⁻¹ compared to the reported affinity of 4-methylcatechol.⁸ This was in agreement with the experimentally determined K_D measured by NMR (see below). FM simulations lead to a ΔG_b° value of -5.5 ± 0.9 kcal mol⁻¹ and a K_D value of 0.11 ± 0.09 mM.

The improved binding of 4-*tert*-butylcatechol may be related to the enhanced van der Waals contacts occurring between the 4-*tert*-butyl group and the neighboring patch of Leu116, Ile119, and Phe120 located in a small helix capping the active site as shown in Figure 4, although this improved interaction does not modify the dispersion of the bound states shared in different conformations. This conformational dispersion was detected after geometrical clustering of the bound conformations (Figure S5), and a similar dispersion was also noticed for 4-methylcatechol.⁸ One cluster included the conformation found in a crystal of hPrx5 in complex with 4-*tert*-butylcatechol (PDB entry 4K7O).⁴⁶

Compared to catechol and 4-methylcatechol, 4-*tert*-butylcatechol bound more on the protein surface next to the active site (green cluster in Figure 4 and Figure S5) with a less stable

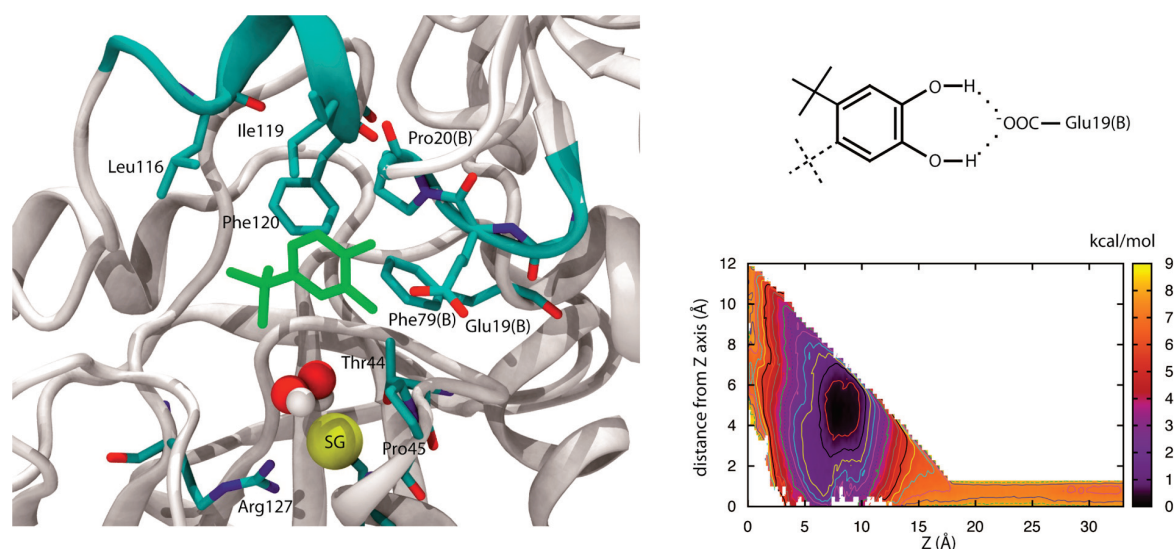


Figure 5. Binding of 4-*tert*-butylcatechol to the hPrx5 active site in the presence of hydrogen peroxide simulated as the Michaelis–Menten complex of hPrx5. The left panel is a representative conformation of the lowest region of the free energy basin. 4-*tert*-Butylcatechol is represented as green sticks and the sulfur of the catalytic Cys47 thiolate as a yellow van der Waals sphere as well as hydrogen peroxide above. Essential amino acids of the active as well as the hydrophobic patch for capping the hPrx5 active site are colored cyan. The top right panel is a schematic of the bound conformation for the ligand's interaction in the lower region of the basin with H-bonds to the Glu19 (chain B) side chain. Alternative conformations are indicated with dashed bonds for the *tert*-butyl group. The bottom right panels is an FES plot represented as a function of the two collective variables used in FM calculations. Contours of iso-energy are drawn on the map using 0.5 kcal mol⁻¹ intervals.

Table 1. Estimation of the Absolute Free Energy Binding Differences (ΔG_b°) and Dissociation Constants (K_D) Using FM for the Free Reduced State (E) of hPrx5 and a Simulated Michaelis–Menten Complex (ES) with Hydrogen Peroxide

	ΔG_b° for E (kcal mol ⁻¹)	ΔG_b° for ES (kcal mol ⁻¹)	K_D for E (mM)	K_D for ES (mM)
catechol	-3.0 ± 0.2^a	-2.0 ± 0.3	6.9 ± 2.1^a	40 ± 30
4-methylcatechol	-4.2 ± 0.3^a	-1.6 ± 0.2	0.9 ± 0.4^a	69 ± 26
4- <i>tert</i> -butylcatechol	-5.5 ± 0.9	-4.4 ± 0.5	0.11 ± 0.09	0.8 ± 0.7

^aValue from ref 8.

conformation observed. This surface interaction is stabilized by a H-bond between the hydroxyl groups of 4-*tert*-butylcatechol and the side-chain carboxylate of Glu19 (chain B) located in a neighboring loop of the active site. Additionally, there were van der Waals contacts that occur with the hydrophobic patch of amino acids (Leu116, Ile119, and Phe120) mentioned above and contacts with Pro20 (chain B) and Phe80 provided by the other side chain of the homodimer (see Figures 4 and 5). The surface conformation is located on the edge of the free energy basin (Figure 4, bottom right), near 8 Å on the Z-axis and approximately 1.5–2 kcal mol⁻¹ above the conformations of the direct interaction with the catalytic Cys47 thiolate. This sort of surface binding was also detected in unbiased NVT trajectories conducted prior to FM. The FES plot features the free energy basin with the angles formed between the catalytic Cys47 sulfur and the oxygen atoms of the hydroxyl groups of 4-*tert*-butylcatechol. The angles are $\sim 180^\circ$ for conformation A and $\sim 60^\circ$ for conformation B (Figure 4, top right).

FM of Catechol Derivatives with a Simulated ES Complex. Determining the relative affinities for the reduced free state of hPrx5 and for the ES complex would improve our understanding and characterization of the type of inhibition mechanism.⁴ The ES complex was modeled by instilling a wall potential limiting the interatomic distance between the sulfur atom of the catalytic Cys47 thiolate and one of the two protons of H₂O₂. The distance of 2.0 Å was retained from the analysis

of the bound states of the substrate in the unbiased NVT MD (see above). No other restraints were used, allowing hydrogen peroxide to adapt continuously to the active site of hPrx5 and the solvent dynamics. The catechol derivatives were then subjected to FM runs and analyzed (Figures S6–S12).

The catechol derivatives were found to have alternative binding modes in the presence of H₂O₂. However, all the catechol derivatives acquired a free energy penalty to form H-bonds with the catalytic thiolate. Figure 5 exemplifies the case of 4-*tert*-butylcatechol with a shift observed in the minimum of the free energy basin. A similar phenomenon was noticed for the two other catechol derivatives given in the Supporting Information (Figures S6–S10). For 4-*tert*-butylcatechol, the major bound conformation fell into the basin corresponding to the surface interaction mentioned above for the free enzyme (see the preceding paragraph and Figure 4).

Table 1 gives the estimated ΔG_b° for each catechol derivative for both the free enzyme and the simulated ES complex. In comparing of the absolute free energy values, the catechol derivatives have a greater affinity for the reduced free state hPrx5 than for the ES complex. Although the catechol derivatives still had an affinity for the ES complex and can still be measured using FM. Therefore, the ability of the catechol derivatives to bind to hPrx5 in a mixed state could suggest the inhibition models against hPrx5.

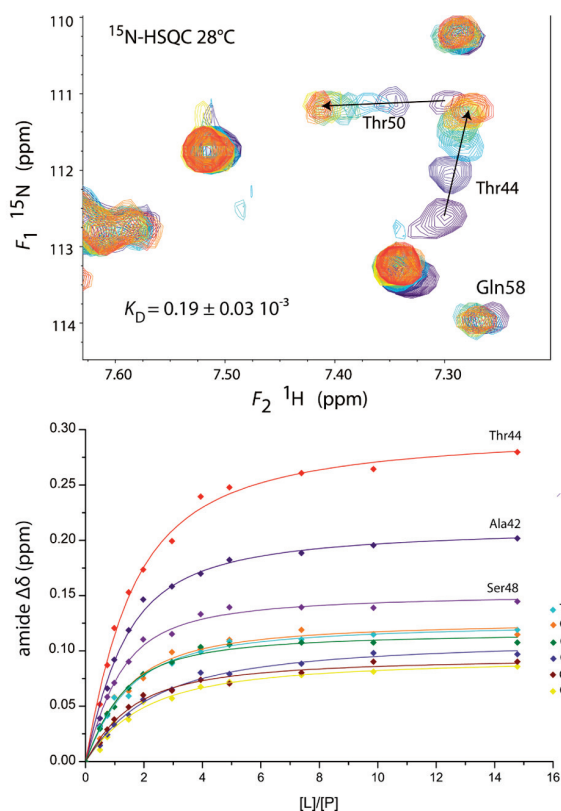


Figure 6. NMR assay of the affinity of 4-*tert*-butylcatechol for reduced hPrx5. The top panel shows superimposed expanded regions of the ^{15}N HSQC experiment, including the Thr50, Thr44, and Gln58 amide signals for protein/ligand ratios ($[\text{L}]/[\text{P}]$) ranging from 0 (blue) to 15 (red). The experiments were conducted at a ^1H frequency of 600 MHz in PBS (pH 7.4) at 28 °C. The arrows indicate the CSPs induced at equilibrium by ligand binding. The bottom panel is a plot of the combined $\Delta\delta$ amide ^1H and ^{15}N CSPs (see Materials and Methods) vs $[\text{L}]/[\text{P}]$.

High-Precision NMR Affinity Assay of 4-*tert*-Butylcatechol for Reduced hPrx5. The binding affinity of 4-*tert*-butylcatechol for hPrx5 was examined by ^{15}N HSQC titration over increasing ligand concentrations. Via assessment of the spectra, the amino acids of hPrx5 having the most significant chemical shift perturbations (CSPs) were distinguished. As in the cases of catechol and 4-methylcatechol,⁸ Ala42, Thr44, and Ser48 contribute to the greatest CSP perturbations for 4-*tert*-butylcatechol (Figure S13). The CSPs were plotted against the ratio concentrations of 4-*tert*-butylcatechol to hPrx5 and were fitted to a hyperbolic isotherm, and a binding dissociation constant (K_D) was obtained (Figure 6). 4-*tert*-Butylcatechol bound to hPrx5 with a K_D of 0.19 ± 0.03 mM, which correlated to the K_D of 0.11 ± 0.09 mM estimated by FM (Table 1). In comparison to all of the catechol derivatives, 4-*tert*-butylcatechol bound to hPrx5 with a binding affinity greater than that of 4-methylcatechol ($K_D = 1.00 \pm 0.20$ mM) or that of catechol ($K_D = 4.50 \pm 0.60$ mM).

IC₅₀ Determination. The ability of the catechol derivatives to bind and inhibit hPrx5 enzymatic activity was first analyzed. hPrx5 enzymatic activity was evaluated through hPrx5 reducing H_2O_2 by NADPH with thioredoxin (γTrxI) and thioredoxin reductase (γTrxR1) as electron transporters (Figure 7). γTrxI , NADPH, and γTrxR1 were used at high concentrations,

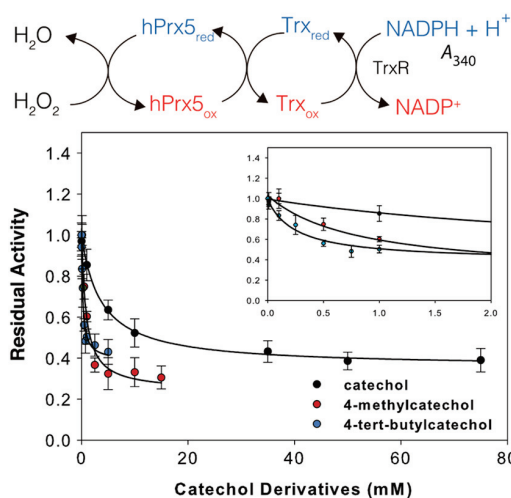


Figure 7. Enzymatic coupling assay used to measure hPrx5 activity and IC_{50} . The top panel shows hPrx5 activity is coupled to NADPH oxidation monitored at 340 nm through thioredoxin (Trx) and thioredoxin reductase (TrxR). The bottom panel is a plot for the concentration-dependent inhibition of hPrx5 by catechol derivatives (Table 2). The reactions were conducted at 23 °C with hPrx5 (0.15 μM) γTrxI (15 μM), γTrxR1 (2 μM), H_2O_2 (25 μM), and NADPH (200 μM).

allowing hPrx5 to be the rate-limiting enzyme.⁴⁷ hPrx5 activity was measured using a substrate concentration around the Michaelis–Menten constant (K_M) value (e.g., ~ 25 μM). The IC_{50} value was determined for each catechol derivative, and the potencies of the inhibitors were established by ranking the IC_{50} values. Under these experimental conditions, 4-*tert*-butylcatechol was determined to be the most potent inhibitor with an IC_{50} value of 0.25 ± 0.06 mM and catechol was the least potent inhibitor with an IC_{50} value of 3.73 ± 0.85 mM (Table 2). Nevertheless, even at high concentrations of the catechol derivatives, the catalytic activity was not fully abolished, suggesting hPrx5 is inhibited through a partial inhibition mechanism (see below).

The possibility of any side reactions or nonspecific interactions of the catechol derivatives with γTrxI , γTrxR1 , or NADPH was considered. Using two-dimensional (^1H vs ^{15}N) NMR, ^{15}N isotope-labeled γTrxI and γTrxR1 chemical shift peaks were monitored before and after their incubation with high concentrations (2–10 mM) of the catechol derivatives. Both γTrxI and γTrxR1 were determined not to bind to any of the catechol derivatives (Figure S14). Also, one-dimensional (1D) ^1H NMR was used to demonstrate that catechol derivatives were not transformed into other by-products that could possibly interfere with any of the reaction components during the enzymatic reaction. For each of the catechol derivatives, the only transformation detected after completion of the enzymes reactions was the full oxidation of NADPH into NADP^+ identified in the 1D spectra (Figure S15).

Type of Inhibition. In identifying how the catechol derivatives could bind and inhibit hPrx5, we examined the inhibition mechanism using steady-state kinetics. As in the IC_{50} assays, hPrx5 activity was monitored using five concentrations of the inhibitors (above and below the IC_{50} value) against a range of H_2O_2 concentrations (0–500 μM).

Table 2. Binding Affinities of Catechol Derivatives and Inhibition Properties of hPrx5^a

	catechol	4-methylcatechol	4- <i>tert</i> -butylcatechol
NMR ¹⁵ N HSQC			
K _D (mM)	4.50 ± 0.60 ^b	1.00 ± 0.20 ^b	0.19 ± 0.03
IC ₅₀			
IC ₅₀ (mM)	3.73 ± 0.85	0.82 ± 0.14	0.25 ± 0.06
DynaFit Model Discrimination (Partial Being the Best Model Type)			
K _s (μM)	40.7 ± 4.0	33.7 ± 4.6	36.9 ± 4.4
K _i (mM)	1.81 ± 0.46	0.33 ± 0.13	0.09 ± 0.04
K _i ' (mM)	4.04 ± 1.13	0.36 ± 0.12	0.18 ± 0.10
k _{cat} (s ⁻¹)	3.08 ± 0.08	3.11 ± 0.01	3.05 ± 0.10
k _{cat} ' (s ⁻¹)	~0 ^c	1.07 ± 0.25	2.35 ± 0.18
k _{cat} /K _s (M ⁻¹ s ⁻¹)	(1.14 ± 0.14) × 10 ⁵	(9.24 ± 1.10) × 10 ⁴	(8.25 ± 0.82) × 10 ⁴

^aThe mode of inhibition by the catechol derivatives of hPrx5 was identified as being partial mixed type noncompetitive (partial). Kinetic parameters K_M, k_{cat}, and k_{cat}' and inhibition constants K_i and K_i' are averaged over three independent repeats with the error being the standard error. ^bFrom ref 8. ^cSee Discussion.

To understand the inhibition mechanism, changes in the Michaelis–Menten constant (K_M) and maximal velocity constant (V_{max}) were assessed. By plotting the reaction rates against the concentration of H₂O₂ through Michaelis–Menten (rate vs [H₂O₂]) (Figure 8) and Hanes–Woolf ([H₂O₂]/rate vs [H₂O₂]) plots (Figures S16 and S17A–C), we interpreted the K_M and V_{max} trends (Tables S1 and S2). Plotting using both methods yielded similar K_M and V_{max} values. All three catechol derivatives followed identical trends with increasing concentrations of the inhibitor; the K_M increased while the V_{max} decreased. Therefore, it was determined that the catechol derivatives inhibit hPrx5 in a noncompetitive manner. This inhibition model is represented by an inhibitor binding to the free enzyme and ES (Figure 9). Furthermore, the inhibition constants (K_i and K_i') were determined to provide more details about the subtype of the noncompetitive mechanism. The inhibition mechanisms were analyzed with DynaFit and interpreted under rapid equilibrium-state and steady-state kinetics in the absence of an inhibitor. DynaFit's analysis was validated under both hypotheses. The equilibrium dissociation constant (K_S = k₋₁/k₁) is equivalent to the Michaelis–Menten constant (K_M) under these experimental settings and complied with the relation k_{cat} ≪ k₋₁. Thus, the rapid equilibrium hypothesis was used to determine the inhibition mechanism using DynaFit and nonlinear regression methods.^{42,43}

For each inhibitor, the reaction rate for each concentration was plotted against the concentration range of H₂O₂ and was fitted to several inhibition mechanism types. All the catechol derivatives were determined to best fit as the partial mixed noncompetitive inhibition mechanism described in Figure 9. For this model, the catechol derivatives can bind to both the hPrx5 and the ES complex. Comparing the inhibition constants (Table 2), the catechol derivatives preferred to bind and inhibit hPrx5. In addition, hPrx5 activity was not fully impaired by the catechol derivatives binding as H₂O₂ could be reduced to water. Two catalytic turnover rates could be determined from ES (k_{cat} or k₂ in eq 1) or the ternary ESI complex (k_{cat}') as shown in Table 2.

DISCUSSION

In investigating the binding of the catechol derivatives to hPrx5, FM and solution NMR experiments provided microscopic details to improve our understanding of the interaction conformations. The catechol derivatives were determined to bind and interact with hPrx5. Also, the catechol derivatives

were identified as inhibitors impeding hPrx5 catalytic activity. By examining the outcomes of the various techniques used, we found parallels with FM and the inhibitory *in vitro* assays. Therefore, the possibility of utilizing FM computational modeling to assist in validating ligands as future inhibitors for drug design development is introduced.

Ligand Binding Affinity Correlates with Inhibition. From the ¹⁵N HSQC NMR titration experiments, 4-*tert*-butylcatechol was determined to bind hPrx5 with a K_D of 0.19 mM. The NMR titrations identified the residues perturbed within the active site of hPrx5 and near catalytic Cys47.^{8,46} 4-*tert*-Butylcatechol had a binding affinity for hPrx5 significantly greater than those of the other catechol derivatives (see Table 2). Additionally, the catechol derivatives, with a binding affinity for hPrx5, were also determined to inhibit hPrx5 catalytic activity. The IC₅₀ assay confirmed 4-*tert*-butylcatechol to be the most potent inhibitor. These findings were supported by a previous study in which 4-methylcatechol was determined to be more potent than catechol, 4-methylcatechol and catechol inhibiting hPrx5 activity with IC₅₀ values of 26 and 105 μM, respectively.³⁷

Catechol Derivative Inhibition Mechanism. The mechanism of inhibition of the catechol derivatives by hPrx5 provided further insight into the binding and inhibition properties. All the catechol derivatives were determined to inhibit in a partial mixed type noncompetitive model (Figure 9). In examining of the inhibitory constant values, the K_i' was greater than or equal to K_D, indicating the inhibitors prefer to bind the free enzyme rather than the ES complex. From a partial mixed type noncompetitive inhibition model, hPrx5 catalysis activity was also not totally abolished. There are two pathways for converting hydrogen peroxide represented by k_{cat} and k_{cat}'.

In examining of 4-*tert*-butylcatechol and 4-methylcatechol inhibition properties, their k_{cat}' values are not as reduced in comparison to their k_{cat} values. This suggested that these two inhibitors do not bind to the ES complex very well (Table 2), whereas for catechol, the k_{cat}' value was practically zero. Therefore, this would indicate that the ESI complex is catalytically inactive and the mechanism could be a mixed type noncompetitive mechanism. However, the catechol's inhibition data fit poorly to this precise model (Figure S18). Upon comparison of the inhibitory constants (K_i and K_i'), the turnover rate values (k_{cat} and k_{cat}'), and the Michaelis–Menten constant (K_M), there were discrepancies in the trends found for

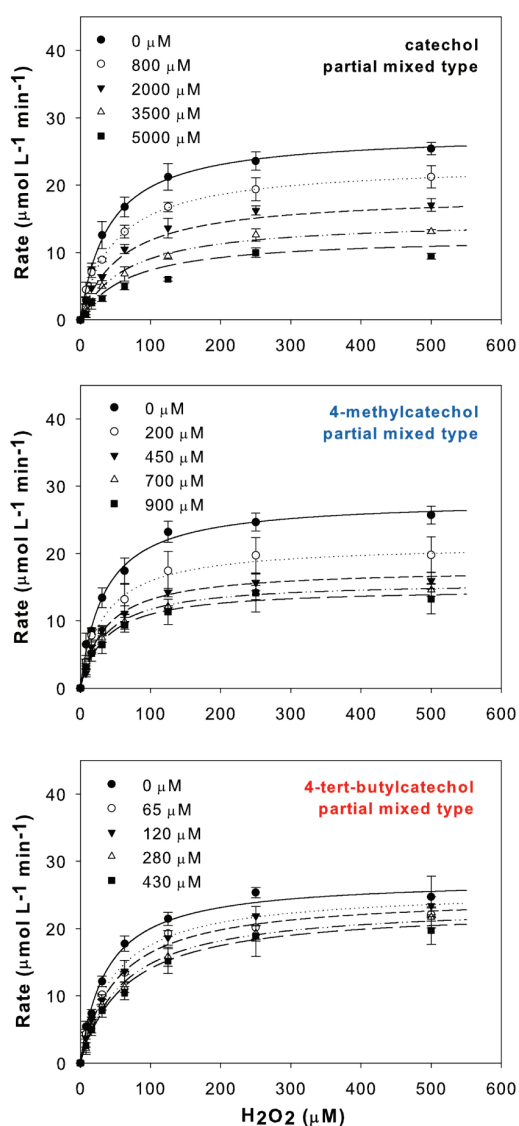


Figure 8. Michaelis–Menten plots of hPrx5 reducing H_2O_2 in the presence of different concentrations of catechol (top), 4-methylcatechol (middle), and 4-*tert*-butylcatechol (bottom). The initial rates for each inhibitor concentration were plotted vs substrate concentration. The data were fitted using a nonlinear fitting method against several types of inhibition models to determine the best inhibition model using Dynafit. The reactions were conducted at 23 °C with hPrx5 (0.15 μM), Trx (15 μM), TrxR (2 μM), H_2O_2 (0–500 μM), NADPH (200 μM), and the catechol derivatives at various concentrations.

the other catechol derivatives. As a result, these differences suggested that catechol does not completely fit to the partial mixed type noncompetitive model but is the closest model.

FM Simulating hPrx5 ES Formation. FM simulations provided valuable details about the reversible binding interactions of H_2O_2 prior to its catalytic conversion. Also, the knowledge of the hPrx5 ES complex is important because the catechol derivatives were determined to inhibit in a noncompetitive manner, affecting the free hPrx5 and ES complex. FM as well as NVT MD trajectories identified a very loose ES complex. H_2O_2 appears to be a very small substrate that diffuses very quickly in solution with no particular affinity

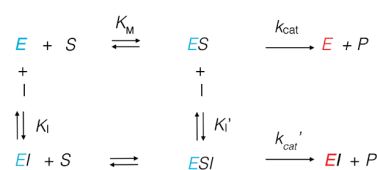


Figure 9. Partial mixed type noncompetitive inhibition model for the catechol derivatives against hPrx5. When hPrx5 (blue) (E) is inhibited by an inhibitor (I), there are two pathways. The I can bind to either the free enzyme or the Michaelis–Menten complex (ES). Under these conditions, hPrx5 is not fully inhibited and can produce product (P): either by ES or ESI represented by rate constant k_{cat} or k_{cat}' , respectively. The oxidized hPrx5 state is colored red (Cys47-SOH or sulfenic acid) and can be regenerated to its reduced state by the thioredoxin system (see Figure 7).

for hPrx5 during the trajectory. The unrestrained NVT trajectories indicated only short residence times on the order of 10^{-9} s. During the binding periods, hydrogen peroxide made a H-bond with the catalytic thiolate and formed a weakly stabilized ES complex. When FM simulations are manipulated, it should be noted that FM increases the number of binding events but the protein dynamics remain unchanged. As a result, this reveals the limitation from FM, and there could be some biased effects in estimating the $\Delta G_{\text{b}}^{\circ}$, for instance, if the hPrx5 needs a longer time scale to transit to another conformation to bind H_2O_2 (a different unknown ES complex).

MD with Ligands. In the unbiased NVT molecular dynamic studies, interaction of hPrx5 with the ligands revealed the same various binding conformations as in FM. This indicated that the external binding potential reproduced the most important interactions well within the active site region. Previously, evaluating the interaction of catechol and 4-methylcatechol with hPrx5 showed that the active site and the whole homodimer are subjected to small conformational changes upon ligand binding. In particular, the 4-methylcatechol active site was observed to induce a small shrinkage of the active site to optimize the interaction with the enzyme.⁸ However, in this case when 4-*tert*-butylcatechol bound, it caused the active site to slightly expand instead of retracting. The active site expansion is probably contributed to the 4-*tert*-butylcatechol being a bulkier ligand and the active site cavity adapting to its size to bind. As also observed in the MD trajectories, the ligands could transit from one bound conformation to another without the need to fully unbind from the active site. These possible transitions are likely related to the fact that the three catechol derivatives still represent low-affinity ligands with K_{D} values ranging from $\sim 10^{-3}$ to 10^{-4} M. Overall, MD and FM simulations contributed to a general understanding of how the catechol derivatives interact and bind to hPrx5. These microscopic aspects of protein–ligand interaction could be advantageous for the production of ideal ligands to become inhibitors of enzymes.

Combining FM and Enzyme Inhibition. Studying the inhibition mechanism against hPrx5, we determined the catechol derivatives inhibited binding to the E and ES complex. FM modeling simulations examined the binding of the catechol derivatives to the E and ES complex. In analyzing the ES complex, we simulated a proposed ES conformation by maintaining hydrogen peroxide bound to hPrx5 through a H-bond to Cys47 thiolate using a distance limiting potential. For each catechol derivative interacting with the mimicked ES complex, there was a drop in the $\Delta G_{\text{b}}^{\circ}$ values of 1–2 kcal

mol⁻¹. Accompanying the changes in the ΔG_b° were alternative modes of binding of the catechol derivatives to the ES complex. For 4-methylcatechol and 4-*tert*-butylcatechols (see Figures 4 and 5), there were additional binding modes mainly seen at the periphery of the hPrx5 active site. These additional binding modes could be attributed to the increased residual affinity of 4-methylcatechol and 4-*tert*-butylcatechol when simulating ES from the extra aliphatic groups present on the catechol backbone. By monitoring the FM simulation, we found 4-methylcatechol and 4-*tert*-butylcatechol interacted with ES but not specifically with Cys47 forming a H-bond. This would be the origin of the significance of the k_{cat}' values noticed for 4-methylcatechol and 4-*tert*-butylcatechol for the inhibition mechanism. However, for catechol, these alternative binding modes were less favorable upon interaction with ES. Catechol's reduced affinity for ES could be associated with the simple structural features of catechols and the inability to make interactions stabilized with Leu116, Ile199, and Phe120 mentioned in Figures 4 and 5.

4-*tert*-Butylcatechol improved binding interactions, which can be attributed to its aliphatic group enhancing noncovalent interactions with the hPrx5 free enzyme state. Also, for interactions of 4-*tert*-butylcatechol with ES, the ligand had alternative binding modes and could complement the retrace-ment in the active site assisting in binding more efficiently. Although these catechol derivatives bound to hPrx5 with a low binding affinity, these microscopic features could be used to improve the design of better ligands to bind to hPrx5. Overall, FM simulations are sufficiently sensitive to measure binding of a ligand to E and ES to discriminate different binding capabilities.

Concluding Remarks. Utilizing NMR techniques, there was early evidence supporting the hypothesis that the catechol derivatives could bind specifically to the active site and do not interact with other catalytic proteins of hPrx5. Accompanied by the NMR results, the FM simulations of the free reduced enzyme and ES complex provided more binding interaction explanations at the microscopic level. Therefore, one could make an inference that FM simulations could be used to predict the inhibition mechanism of a ligand. The estimated ΔG_b° for both the free enzyme and a simulated ES from FM had common a trend with respect to the inhibition constants (K_i and K_i'). Consequently, one could anticipate the inhibition model as being a noncompetitive mechanism specifically for the catechol derivatives against hPrx5.

Here this study shows how integrating solution NMR and MD simulations can be a valuable tool in understanding the protein–ligand interaction at the microscopic level. Specifically, MD analysis can contribute to the details used for preliminary predictions of potential ligands being inhibitors and the proposal of an inhibition mechanism. This is one example of how MD simulations can be used in drug design to improve drug development. However, additional studies using MD to predict ligands as potential inhibitors will strengthen and improve the value of computational simulations. Overall, MD can contribute to improving drug screening time and costs, benefiting the *in vitro* investigation level in the pharmaceutical industry.

■ ASSOCIATED CONTENT

● Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.biochem.6b00367.

Figures S1–S17 and two tables providing information about MD and enzyme inhibition kinetics (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: jean-marc.lancelin@isa-lyon.fr.

Funding

M.L.C. is a recipient of a Ph.D. grant 2013–2016 to promote the international mobility of Université Claude Bernard Lyon 1. Funding of this project was provided by a grant from Région Rhône-Alpes, France, N° 13-0118962-01 to L.T. and J.-M.L.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We are grateful to Prof. Bernard Knoops and André Clippe from University of Louvain-la-Neuve for the generous gifts of biological materials for expressing and purifying the recombinant proteins. We thank Acellera Co. for providing a complementary 6-month ACEMD multi-GPU software license and for the generous gift of a Nvidia Tesla K20 supercomputing unit.

■ ABBREVIATIONS

CSP, chemical shift perturbation; FM, funnel metadynamics; MD, molecular dynamics; hPrx5, human peroxiredoxin; yTrx1, yeast thioredoxin 1; yTrxR1, yeast thioredoxin reductase 1

■ REFERENCES

- (1) Henri, V. (1902) Théorie générale de l'action de quelques diastases. *C. R. Acad. Sci. Paris* 135, 916–919.
- (2) Michaelis, L., and Menten, M. L. (1913) Kinetik der Invertinwirkung. *Biochem. Zeitung* 49, 333–369.
- (3) Johnson, K. A., and Goody, R. S. (2011) The Original Michaelis Constant: Translation of the 1913 Michaelis–Menten Paper. *Biochemistry* 50, 8264–8269.
- (4) Segel, I. (1993) *Enzymes Kinetics. Behavior Analysis of Rapid Equilibrium and Steady-State Enzymes Systems*, Wiley Classics Library Edition, Wiley, New York.
- (5) Cornish-Bowden, A. (2015) One hundred years of Michaelis–Menten kinetics. *Persp. Sci.* 4, 3–9.
- (6) Changeux, J. P. (2011) 50th anniversary of the word "allosteric". *Prot. Sci.* 20, 1119–1124.
- (7) Harvey, M. J., and De Fabritiis, G. (2012) High-throughput molecular dynamics: the powerful new tool for drug discovery. *Drug Discovery Today* 17, 1059–1062.
- (8) Troussicot, L., Guilliere, F., Limongelli, V., Walker, O., and Lancelin, J. M. (2015) Funnel-metadynamics and solution NMR to estimate protein-ligand affinities. *J. Am. Chem. Soc.* 137, 1273–1281.
- (9) Limongelli, V., Bonomi, M., and Parrinello, M. (2013) Funnel metadynamics as accurate binding free-energy method. *Proc. Natl. Acad. Sci. U. S. A.* 110, 6358–6363.
- (10) Karplus, P. A. (2015) A primer on peroxiredoxin biochemistry. *Free Radical Biol. Med.* 80, 183–190.
- (11) Perkins, A., Nelson, K. J., Parsonage, D., Poole, L. B., and Karplus, P. A. (2015) Peroxiredoxins: guardians against oxidative stress and modulators of peroxide signaling. *Trends Biochem. Sci.* 40, 435–445.

- (12) Wood, Z. A., Schroder, E., Robin Harris, J., and Poole, L. B. (2003) Structure, mechanism and regulation of peroxiredoxins. *Trends Biochem. Sci.* 28, 32–40.
- (13) Wood, Z. A., Poole, L. B., and Karplus, P. A. (2003) Peroxiredoxin evolution and the regulation of hydrogen peroxide signaling. *Science* 300, 650–653.
- (14) Seo, M. S., Kang, S. W., Kim, K., Baines, I. C., Lee, T. H., and Rhee, S. G. (2000) Identification of a new type of mammalian peroxiredoxin that forms an intramolecular disulfide as a reaction intermediate. *J. Biol. Chem.* 275, 20346–20354.
- (15) Perkins, A., Poole, L. B., and Karplus, P. A. (2014) Tuning of Peroxiredoxin Catalysis for Various Physiological Roles. *Biochemistry* 53, 7693–7705.
- (16) Fujii, J., and Ikeda, Y. (2002) Advances in our understanding of peroxiredoxin, a multifunctional, mammalian redox protein. *Redox Rep.* 7, 123–130.
- (17) Nystrom, T., Yang, J., and Molin, M. (2012) Peroxiredoxins, gerontogenes linking aging to genome instability and cancer. *Genes Dev.* 26, 2001–2008.
- (18) Shichita, T., Hasegawa, E., Kimura, A., Morita, R., Sakaguchi, R., Takada, I., Sekiya, T., Ooboshi, H., Kitazono, T., Yanagawa, T., Ishii, T., Takahashi, H., Mori, S., Nishibori, M., Kuroda, K., Akira, S., Miyake, K., and Yoshimura, A. (2012) Peroxiredoxin family proteins are key initiators of post-ischemic inflammation in the brain. *Nat. Med.* 18, 911–917.
- (19) Haraldsen, J. D., Liu, G., Botting, C. H., Walton, J. G. A., Storm, J., Phalen, T. J., Kwok, L. Y., Soldati-Favre, D., Heintz, N. H., Muller, S., Westwood, N. J., and Ward, G. E. (2009) Identification of conoidin A as a covalent inhibitor of peroxiredoxin II. *Org. Biomol. Chem.* 7, 3040–3048.
- (20) Liu, C. X., Yin, Q. Q., Zhou, H. C., Wu, Y. L., Pu, J. X., Xia, L., Liu, W., Huang, X., Jiang, T., Wu, M. X., He, L. C., Zhao, Y. X., Wang, X. L., Xiao, W. L., Chen, H. Z., Zhao, Q., Zhou, A. W., Wang, L. S., Sun, H. D., and Chen, G. Q. (2012) Adenanthin targets peroxiredoxin I and II to induce differentiation of leukemic cells. *Nat. Chem. Biol.* 8, 486–493.
- (21) Wei, W., Ma, C., Cao, Y., Yang, L., Huang, Z., Qin, D., Chen, Y., Liu, C., Xia, L., and Wang, T. (2016) Identification of H7 as a novel peroxiredoxin I inhibitor to induce differentiation of leukemia cells. *Oncotarget* 7, 3873–3883.
- (22) Zhao, Q., Ding, Y., Deng, Z., Lee, O.-Y., Gao, P., Chen, P., Rose, R. J., Zhao, H., Zhang, Z., Tao, X.-P., Heck, A. J. R., Kao, R., and Yang, D. (2015) Natural products triptolide, celastrol, and withaferin A inhibit the chaperone activity of peroxiredoxin I. *Chem. Sci.* 6, 4124–4130.
- (23) Parker, A. R., Petluru, P. N., Nienaber, V. L., Badger, J., Leverett, B. D., Jair, K., Sridhar, V., Logan, C., Ayala, P. Y., Kochat, H., and Hausheer, F. H. (2015) Cysteine specific targeting of the functionally distinct peroxiredoxin and glutaredoxin proteins by the investigational disulfide BNP7787. *Molecules* 20, 4928–4950.
- (24) Hall, A., Parsonage, D., Poole, L. B., and Karplus, P. A. (2010) Structural evidence that peroxiredoxin catalytic power is based on transition-state stabilization. *J. Mol. Biol.* 402, 194–209.
- (25) Bernstein, F. C., Koetzle, T. F., Williams, G. J., Meyer, E. F., Jr., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., and Tasumi, M. (1978) The Protein Data Bank: a computer-based archival file for macromolecular structures. *Arch. Biochem. Biophys.* 185, 584–591.
- (26) Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., and Kollman, P. A. (1995) A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* 117, 5179–5197.
- (27) Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., and Simmerling, C. (2006) Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins: Struct., Funct., Genet.* 65, 712–725.
- (28) Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J. L., Dror, R. O., and Shaw, D. E. (2010) Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins: Struct., Funct., Genet.* 78, 1950–1958.
- (29) Jorgensen, W. L., and Madura, J. D. (1983) Quantum and statistical mechanical studies of liquids. 25. Solvation and conformation of methanol in water. *J. Am. Chem. Soc.* 105, 1407–1413.
- (30) Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., and Case, D. A. (2004) Development and testing of a general amber force field. *J. Comput. Chem.* 25, 1157–1174.
- (31) Wang, J., Wang, W., Kollman, P. A., and Case, D. A. (2006) Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graphics Modell.* 25, 247–260.
- (32) Harvey, M. J., Giupponi, G., and Fabritiis, G. D. (2009) ACEMD: Accelerating Biomolecular Dynamics in the Microsecond Time Scale. *J. Chem. Theory Comput.* 5, 1632–1639.
- (33) Feenstra, K. A., Hess, B., and Berendsen, H. J. C. (1999) Improving efficiency of large time-scale molecular dynamics simulations of hydrogen-rich systems. *J. Comput. Chem.* 20, 786–798.
- (34) Bonomi, M., Branduardi, D., Bussi, G., Camilloni, C., Provasi, D., Raiteri, P., Donadio, D., Marinelli, F., Pietrucci, F., Broglia, R. A., and Parrinello, M. (2009) PLUMED: A portable plugin for free-energy calculations with molecular dynamics. *Comput. Phys. Commun.* 180, 1961–1972.
- (35) Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: Visual molecular dynamics. *J. Mol. Graphics* 14, 33–38.
- (36) Declercq, J. P., Evrard, C., Clippe, A., Stricht, D. V., Bernard, A., and Knoops, B. (2001) Crystal structure of human peroxiredoxin 5, a novel type of mammalian peroxiredoxin at 1.5 Å resolution. *J. Mol. Biol.* 311, 751–759.
- (37) Barelier, S., Linard, D., Pons, J., Clippe, A., Knoops, B., Lancelin, J. M., and Krimm, I. (2010) Discovery of fragment molecules that bind the human peroxiredoxin 5 active site. *PLoS One* 5, e9744.
- (38) Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* 6, 277–293.
- (39) Johnson, B., and Blevins, R. (1994) NMR View: A computer program for the visualization and analysis of NMR data. *J. Biomol. NMR* 4, 603–614.
- (40) Farmer, B. T., 2nd, Constantine, K. L., Goldfarb, V., Friedrichs, M. S., Wittekind, M., Yanchunas, J., Jr., Robertson, J. G., and Mueller, L. (1996) Localizing the NADP⁺ binding site on the MurB enzyme by NMR. *Nat. Struct. Biol.* 3, 995–997.
- (41) Kim, I., Lee, K. S., Hwang, J. S., Ahn, M. Y., Li, J., Sohn, H. D., and Jin, B. R. (2005) Molecular cloning and characterization of a peroxiredoxin gene from the mole cricket, *Gryllotalpa orientalis*. *Comp. Biochem. Physiol., Part B: Biochem. Mol. Biol.* 140, 579–587.
- (42) Kuzmič, P. (2009) DynaFit: A Software Package for Enzymology. In *Methods in Enzymology*, Chapter 10, pp 247–280, Academic Press, San Diego.
- (43) Kuzmič, P. (1996) Program DYNAFIT for the Analysis of Enzyme Kinetic Data: Application to HIV Proteinase. *Anal. Biochem.* 237, 260–273.
- (44) Laio, A., and Gervasio, F. L. (2008) Metadynamics: a method to simulate rare events and reconstruct the free energy in biophysics, chemistry and material science. *Rep. Prog. Phys.* 71, 126601.
- (45) Barducci, A., Bussi, G., and Parrinello, M. (2008) Well-Tempered Metadynamics: A Smoothly Converging and Tunable Free-Energy Method. *Phys. Rev. Lett.* 100, 020603.
- (46) Aguirre, C., ten Brink, T., Guichou, J. F., Cala, O., and Krimm, I. (2014) Comparing binding modes of analogous fragments using NMR in fragment-based drug design: application to PRDX5. *PLoS One* 9, e102300.
- (47) Copeland, R. A. (2005) Evaluation of enzyme inhibitors in drug discovery. A guide for medicinal chemists and pharmacologists. *Methods Biochem. Anal.* 46, 1–265.

Predicting and Understanding the Enzymatic Inhibition of Human Peroxiredoxin 5 by 4-Substituted Pyrocatechols Combining Funnel-Metadynamics, Solution NMR and Steady-State Kinetics

Melissa L. Chow[§], Laura Troussicot[§], Marie Martin[§], Bastien Doumèche[¶], Florence Guillière[§], Jean-Marc Lancelin^{§*}

[§] Université de Lyon, Institut des Sciences Analytiques, UMR 5280, CNRS, Université Lyon 1, ENS Lyon - 5, rue de la Doua, F-69100 Villeurbanne, France. [¶] Université de Lyon, Institut de Chimie et Biochimie Moléculaires et Supramoléculaires, UMR 5246, Université Claude Bernard Lyon1, CNRS, CPE-Lyon, 43 Bd du 11 novembre 1918, F-69622, Villeurbanne, France

KEYWORDS: Funnel-metadynamics, enzyme kinetics, enzyme inhibition, solution NMR, human peroxiredoxin 5.

SUPPORTING INFORMATION

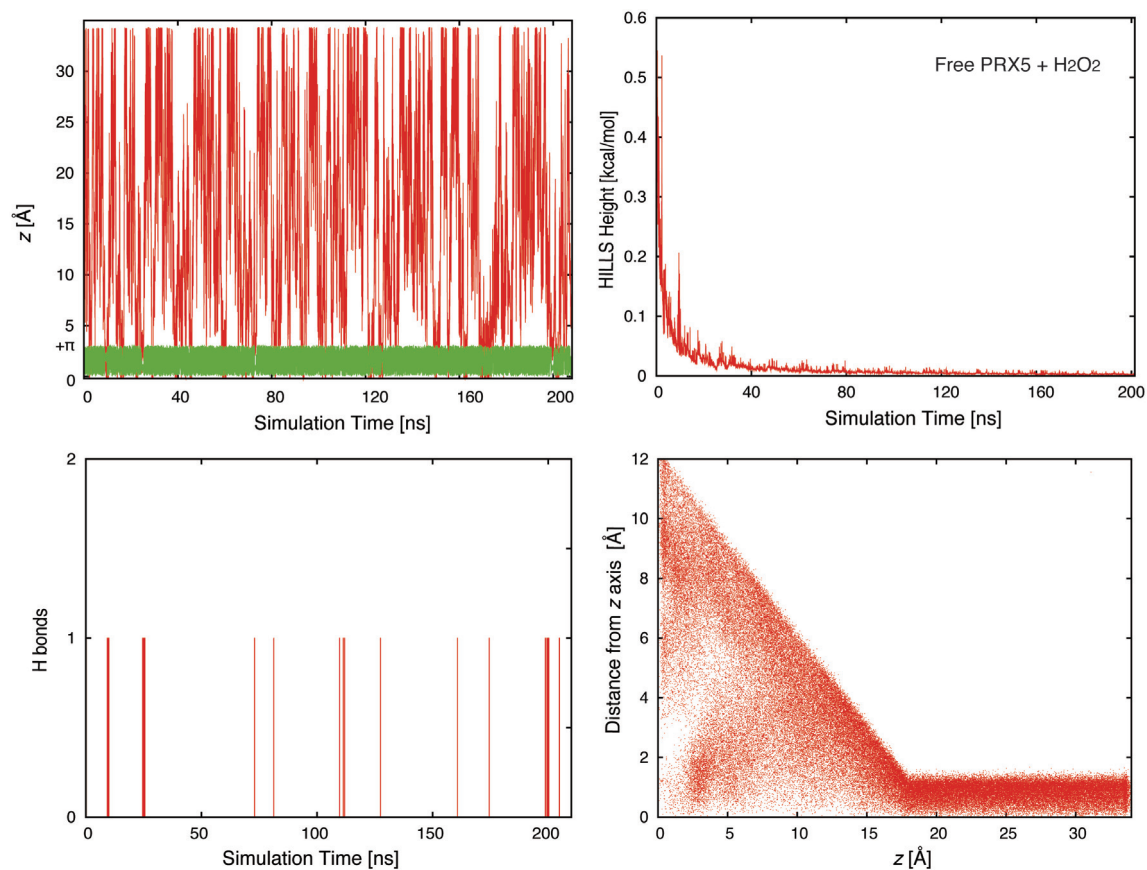


Figure S1. Funnelmetadynamics trajectories of reduced human peroxiredoxin 5 (hPrx5) in the presence of hydrogen peroxide. Upper left panel, red trace : the position of the center of mass for the two oxygen atoms of hydrogen peroxide projected on the funnel Z -axis in Å (CV1) as a function of the simulation time. The ligand samples the whole Z -axis dimension with the defined 0-34 Å limits of the fixed funnel external potential. The green trace indicate the collective variable Cys47-S⁻(PRX5)...O1-O2 (hydroperoxide) angle (CV3) reported in radians. Above Z -axis=18 Å. The ligand samples all conformations in the cylinder section of the funnel and interacts only with the solvent. Below 18 Å, the ligand interacts with the protein in the region of the active site. The binding periods correspond to a defined CV3 and a short CV1. Lower left panel : hydrogen bond between Cys47 and hydrogen peroxide count vs. FM simulation time indicating the binding periods. Upper-right panel : Hills potentials deposition in kcal mol⁻¹ along the simulation time indicating the progressing convergence where no free-energy depositions occur anymore. Lower-right panel : hydrogen peroxide trajectories according to collective variables CV1 and CV2 (positional distance of the ligand to the funnel Z -axis).

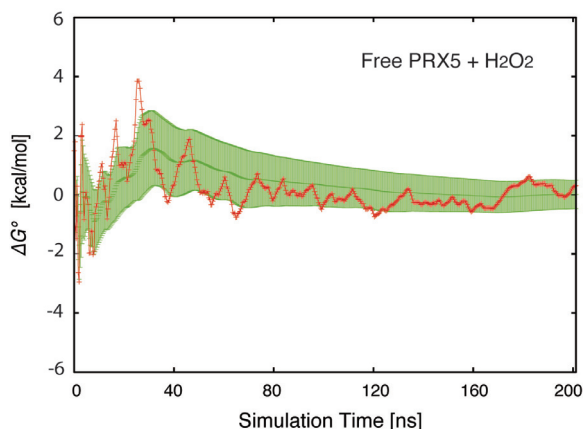


Figure S2. Evolution of the calculated absolute binding free-energy (ΔG_b^0) between reduced hPrx5 (red traces and points) and its substrate hydrogen peroxide as function of the funnel-metadynamics simulation time. The value is a weighted average (green trace) and the error represented as standard deviation (green bars). ΔG_b^0 converges to a value near zero indicating a weakly bound enzyme-substrate complex before catalysis.

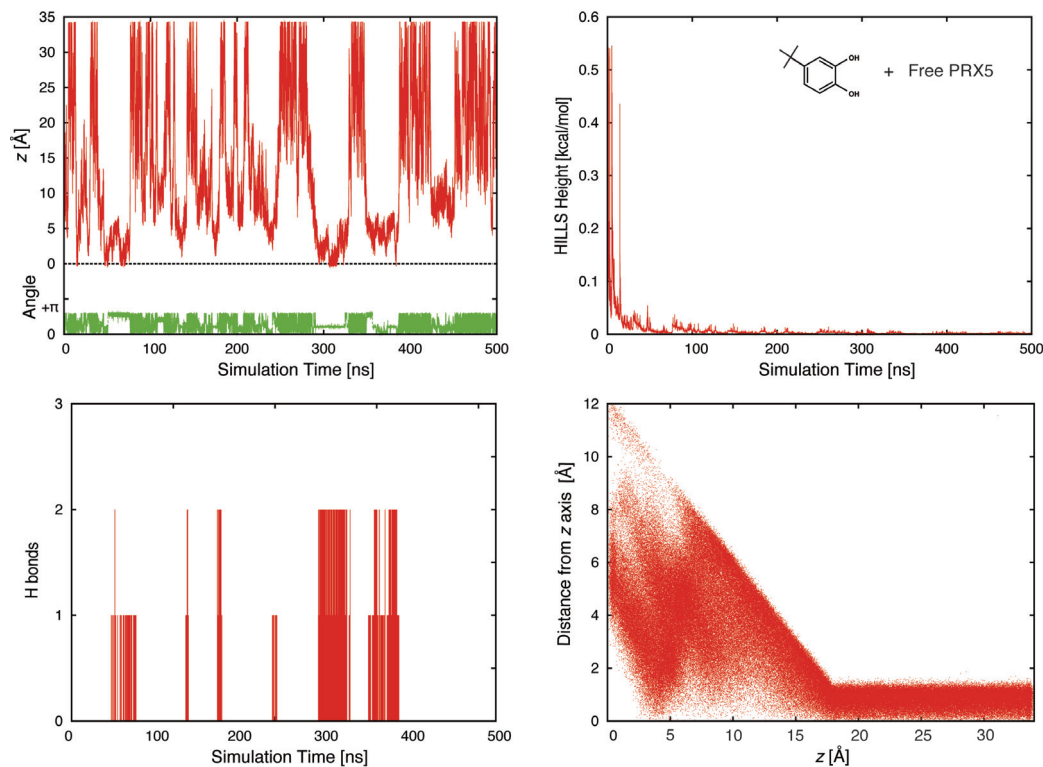


Figure S3. Funnel-metadynamics trajectories of reduced hPx5 in the presence of 4-tert-butylcatechol. Upper left panel, red trace : position of the center of mass of the ligand projected on the funnel Z-axis in Å (CV1) as a function of the simulation time. The green trace indicates the collective variable Cys47-S (PRX5) ... O1-O2 (4-tert-butylcatechol) angle (CV3) reported in radian. The binding periods correspond to a defined CV3 and a short CV1. Lower left panel : hydrogen bond between Cys47 and 4-tert-butyl catechol counts vs. FM simulation time indicating the binding periods. Binding periods are longer than for the low-affinity enzyme substrate due to the higher affinity of the ligands (inhibitors). Upper-right panel : Hills potentials deposition in kcal mol⁻¹ along the simulation time indicating the progressing convergence where no free-energy depositions occur no more. Lower-right panel : 4-tert-butylcatechol trajectories according to collective variables CV1 and CV2 (positional distance of the ligand to the funnel Z-axis).

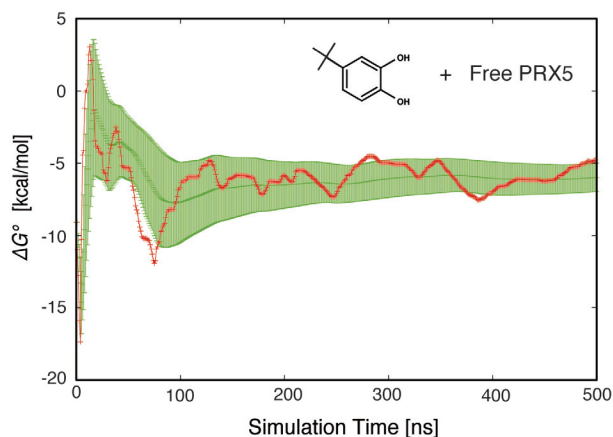


Figure S4. Calculated absolute binding free-energy (ΔG_b^0) between reduced hPrx5 (red traces and points) and 4-tert-butylcatechol as a function of the funnel-metadynamics simulation time. The ΔG_b^0 is a weighted average (green trace) and with its error shown as the standard deviation (green bars).

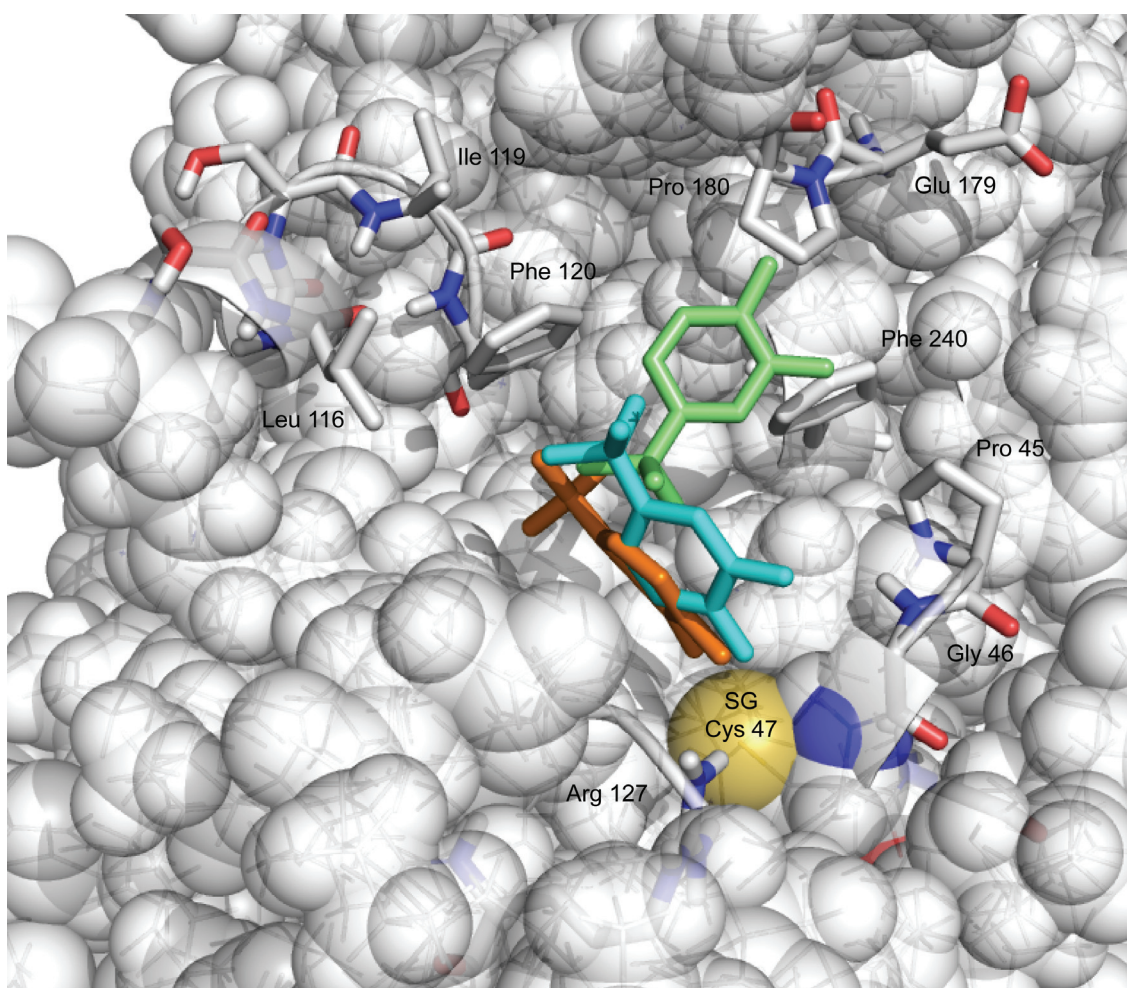


Figure S5 – Geometrical clustering of the funnel-metadynamics trajectories of reduced hPxS in the presence of 4-*tert*-butylcatechol. Geometrical clustering were done using 2500 frames collected every 200 ps of the 500 ns trajectory presented in Fig. S3 and S4. The VMD software plugin tool Clustering V2.0 of Luis Garcia, Cornell Medical College, USA, was used to cluster geometrically the 2500 conformations using a cutoff RMS difference of 2Å. For clarity, the results were reduced to 3 clusters. History-dependent analysis of the trajectory made in VMD indicates that the green conformer is not a pre-requisite conformation to reach the lower basin A () and B () that can be reached independently.

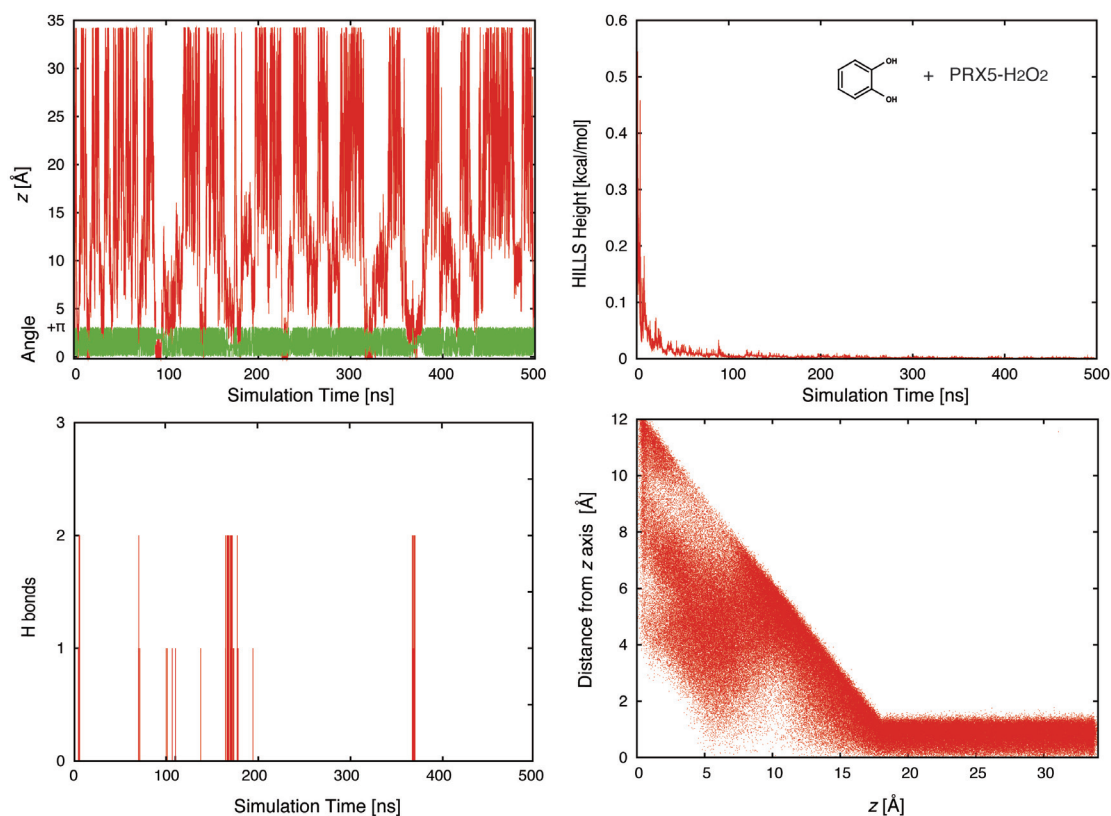


Figure S6. Funnel-metadynamics trajectories of reduced hPrx5 in the presence of catechol and the enzyme substrate hydrogen peroxide maintained at a maximum Cys47-S' (PRX5)... H-O1 (hydroperoxide) interatomic distance of 2 Å with a wall potential of 20 kcal.mol⁻¹.Å⁻². Upper left panel, red trace : the position of the center of mass catechol projected on the funnel Z-axis in Å (CV1) as a function during the simulation time. The green trace indicates the collective variable Cys47-S' (PRX5)... O1-O2 (catechol) angle (CV3) reported in radian. The binding periods correspond to a defined CV3 and a short CV1. Lower left panel : hydrogen bond between Cys47 and catechol counts vs. FM simulation time indicating the binding periods. Upper-right panel : Hills potentials deposition in kcal.mol⁻¹ along the simulation time indicating the progressing convergence where no energy depositions occurs no more. Lower-right panel : catechol trajectories according to collective variables CV1 and CV2 (positional distance of the ligand to the funnel Z-axis).

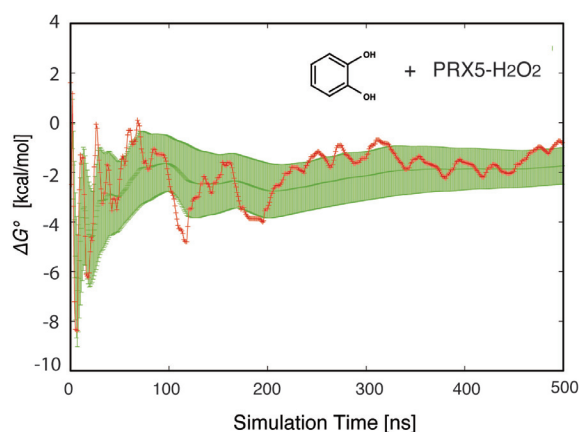


Figure S7. Evolution of the calculated absolute binding free energy (ΔG_b^0) between hPrx5 (red traces and points) in the presence of catechol and the enzyme substrate hydrogen peroxide maintained at a maximum Cys47-S' (PRX5)... H-O1 (hydroperoxide) interatomic distance of 2 Å with a wall potential of 20 kcal.mol⁻¹, as function of the funnel-metadynamics simulation time. The ΔG_b^0 is a weighted average (green trace) and with its error shown as standard deviation (green bars).

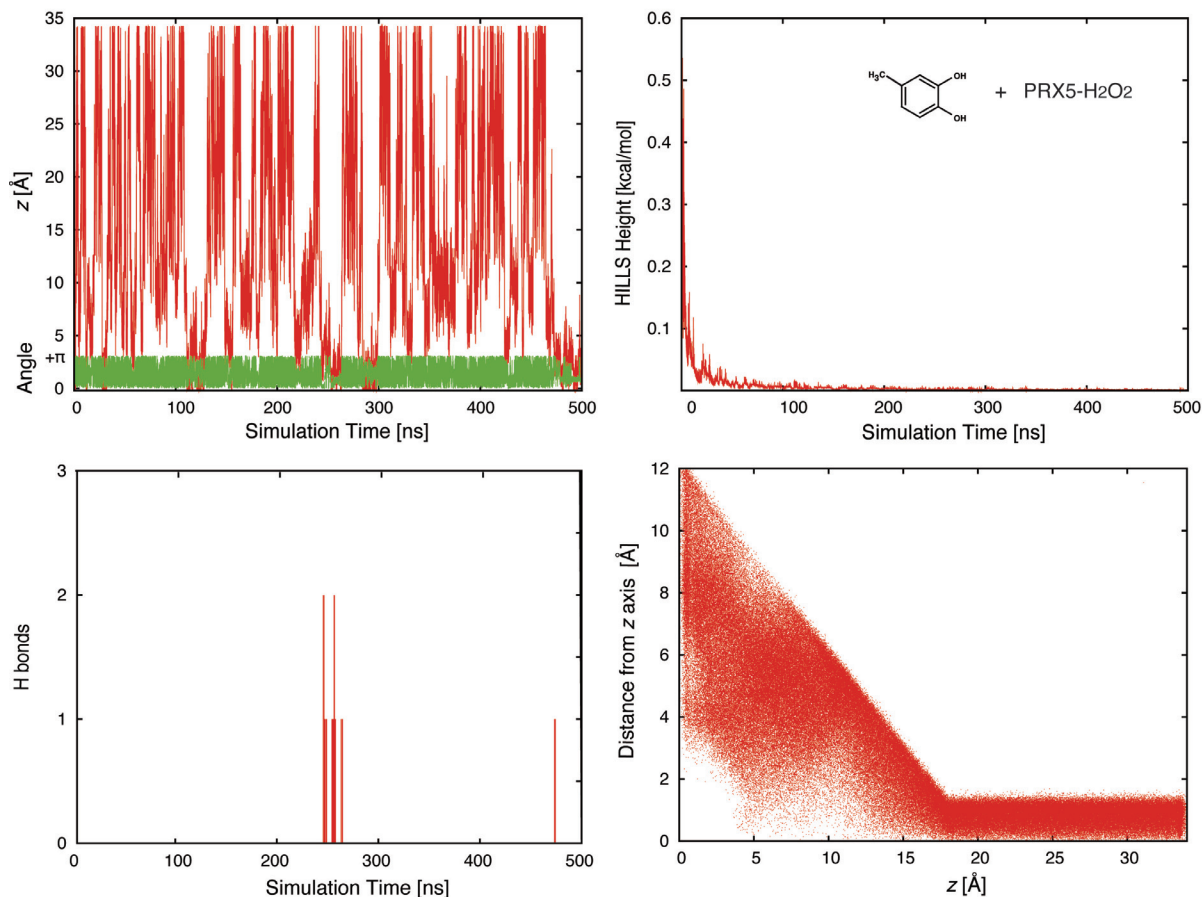


Figure S8. Funnel-metadynamics trajectories of reduced hPx5 in the presence of 4-methylcatechol and the enzyme substrate hydrogen peroxide maintained at a maximum Cys47-S' (PRX5) ... H-O1 (hydroperoxide) interatomic distance of 2 Å with a wall potential of 20 kcal.mol⁻¹.Å⁻². Upper left panel, red trace : the position of the center of mass 4-methylcatechol projected on the funnel Z-axis in Å (CV1) as a function of the simulation time. The green trace indicates the collective variable Cys47-S' (PRX5)...O1-O2 (4-methylcatechol) angle (CV3) reported in radian. The binding periods correspond to a defined CV3 and a short CV1. Lower left panel : hydrogen bond between Cys47 and 4-methylcatechol counts vs. FM simulation time indicating the binding periods. Upper-right panel : Hills potentials deposition in kcal.mol⁻¹ along the simulation time indicating the progressing convergence where no free-energy depositions occur no more. Lower-right panel : 4-methylcatechol trajectories according to collective variables CV1 and CV2 (positional distance of the ligand to the funnel Z-axis).

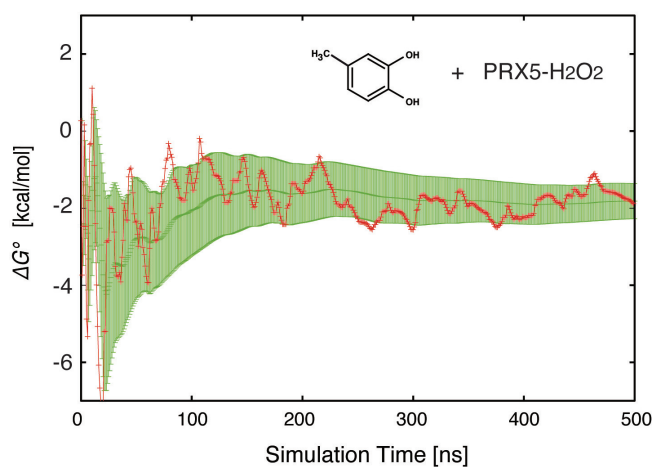


Figure S9. Calculated absolute binding free energy (ΔG_b^0) between reduced hPrx5 (red traces and points) in the presence 4-methylcatechol and the enzyme substrate hydrogen peroxide maintained at a maximum Cys47-S' (PRX5) ... H-O1 (hydroperoxide) interatomic distance of 2 Å with a wall potential of 20 kcal.mol⁻¹, as function of the funnel-metadynamics simulation time. The ΔG_b^0 it's a weighted average (green trace) and with its error shown as standard deviation (green bars).

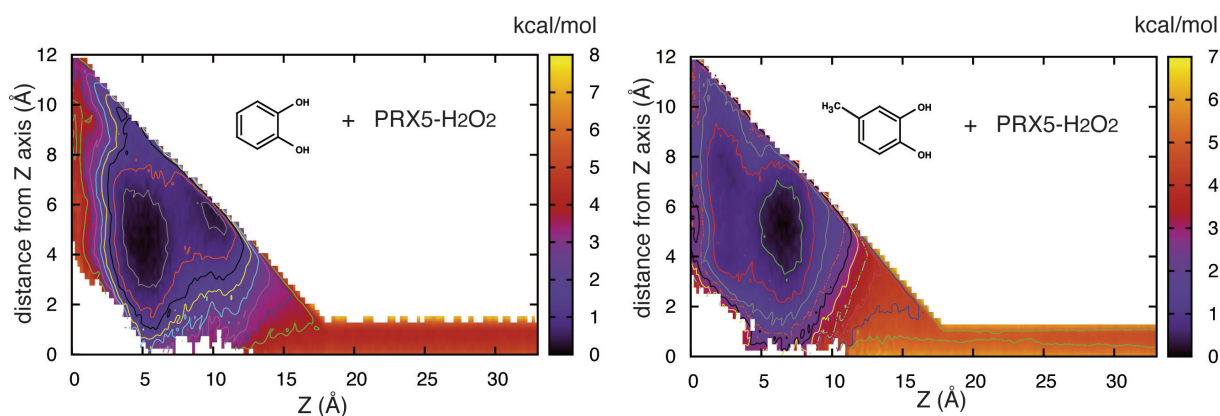


Figure S10. Free-Energy Surfaces (FES) represented as a function of the 2 collective variables used in funnel-metadynamics calculations of reduced hPrx5 in the presence catechol (left panel) or 4-methylcatechol (right panel) and the enzyme substrate hydrogen peroxide maintained at a maximum Cys47-S' (PRX5) ... H-O1 (hydroperoxide) interatomic distance of 2 Å with a wall potential of 20 kcal.mol⁻¹. Contours of iso-energy are drawn on the map using 0.5 kcal.mol⁻¹.

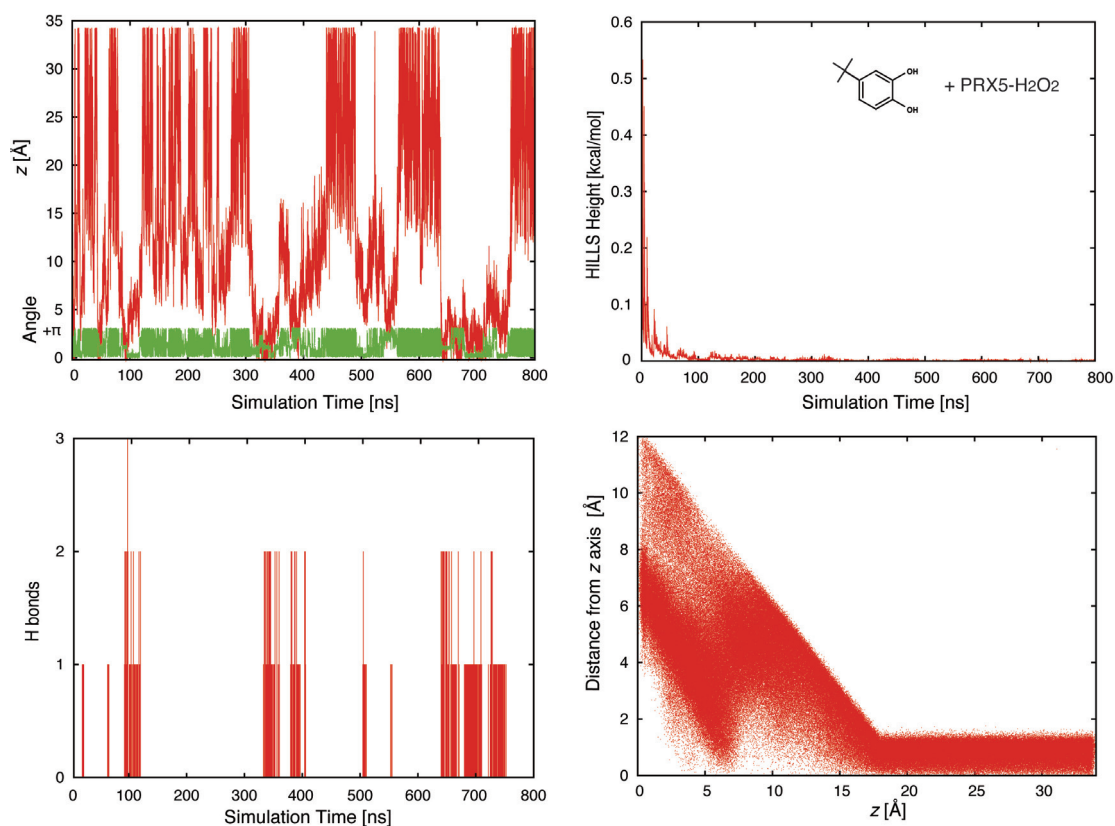


Figure S11. Funnel-metadynamics trajectories of reduced hPrx5 in the presence 4-*tert*-butylcatechol and the enzyme substrate hydrogen peroxide maintained at a maximum Cys47-S' (hPrx5) ... H-O1 (hydrogen peroxide) interatomic distance of 2 Å with a wall potential of 20 kcal.mol⁻¹.Å². Upper left panel, red trace : the position of the center of mass of 4-*tert*-butylcatechol projected on the funnel Z-axis in Å (CV1) as a function of the simulation time. The green trace indicates the collective variable Cys47-S' (PRX5) ... O1-O2 (4-*tert*-butylcatechol) angle (CV3) reported in radian. The binding periods correspond to a defined CV3 and a short CV1. Lower left panel : hydrogen bond between Cys47 and 4-methylcatechol counts vs. FM simulation time indicating the binding periods. Upper-right panel : Hills potentials deposition in kcal.mol⁻¹ along the simulation time indicating the progressing convergence where no free-energy depositions occur no more. Lower-right panel : 4-*tert*-butylcatechol trajectories according to collective variables CV1 and CV2 (positional distance of the ligand to the funnel Z-axis).

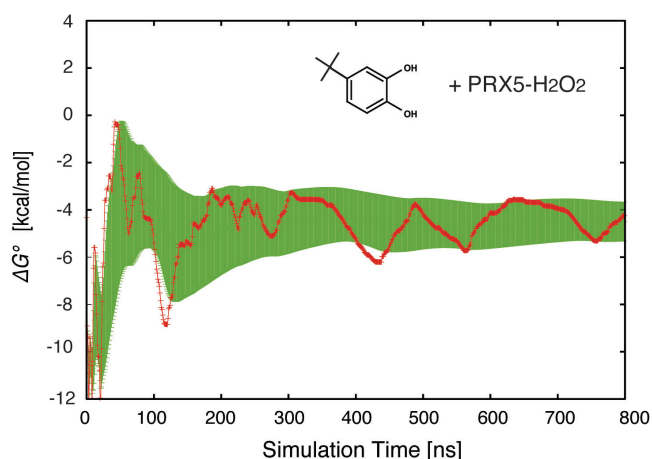


Figure S12. Calculated absolute binding free-energy (ΔG_b^0) between reduced hPrx5 (red traces and points) in the presence 4-*tert*-butylcatechol and the enzyme substrate hydrogen peroxide maintained at a maximum Cys47-S' (PRX5)... H-O1 (hydrogen peroxide) interatomic distance of 2 Å with a wall potential of 20 kcal.mol⁻¹.Å⁻², as function of the funnel-metadynamics simulation time. The ΔG_b^0 it's a weighted average (green trace) and with its error shown as standard deviation (green bars).

Comparison of induced CSPs for [Ligands] ~ 2KD

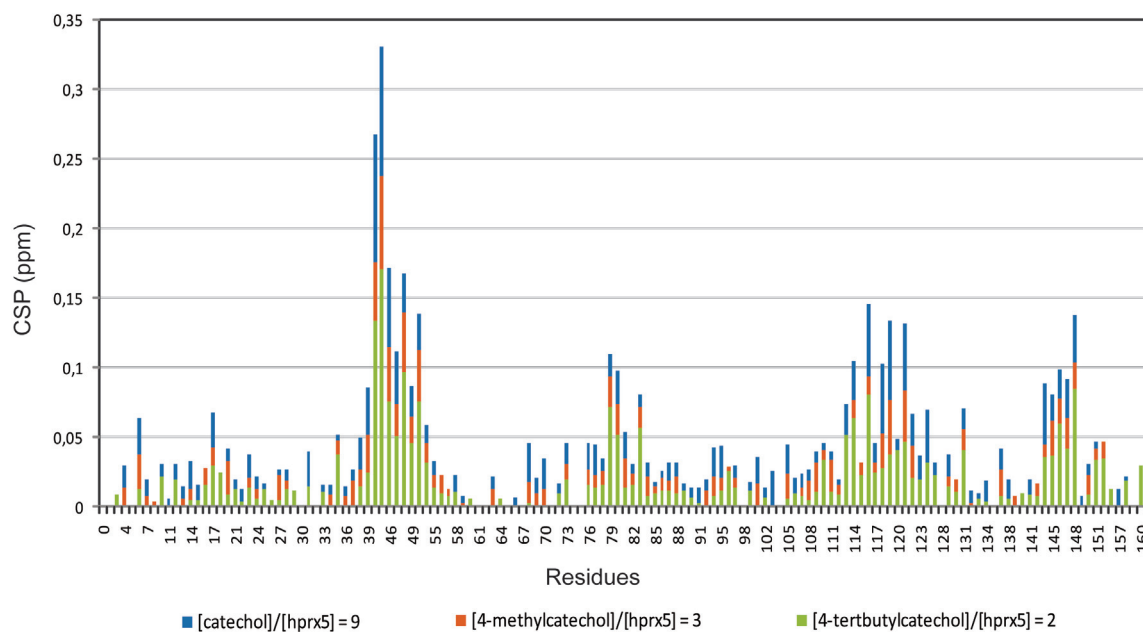


Figure S13. Stacked histogram of chemical shift perturbations (CSPs) induced to hPrx5 ¹H-¹⁵N NMR signals by catechol derivatives : catechol (blue), 4-methylcatechol (orange) and 4-*tert*-butylcatechol (green) at ligands concentrations corresponding to 67% of the hPrx5 saturation for each. a due to spectral or non-observed amino acids are not represented.

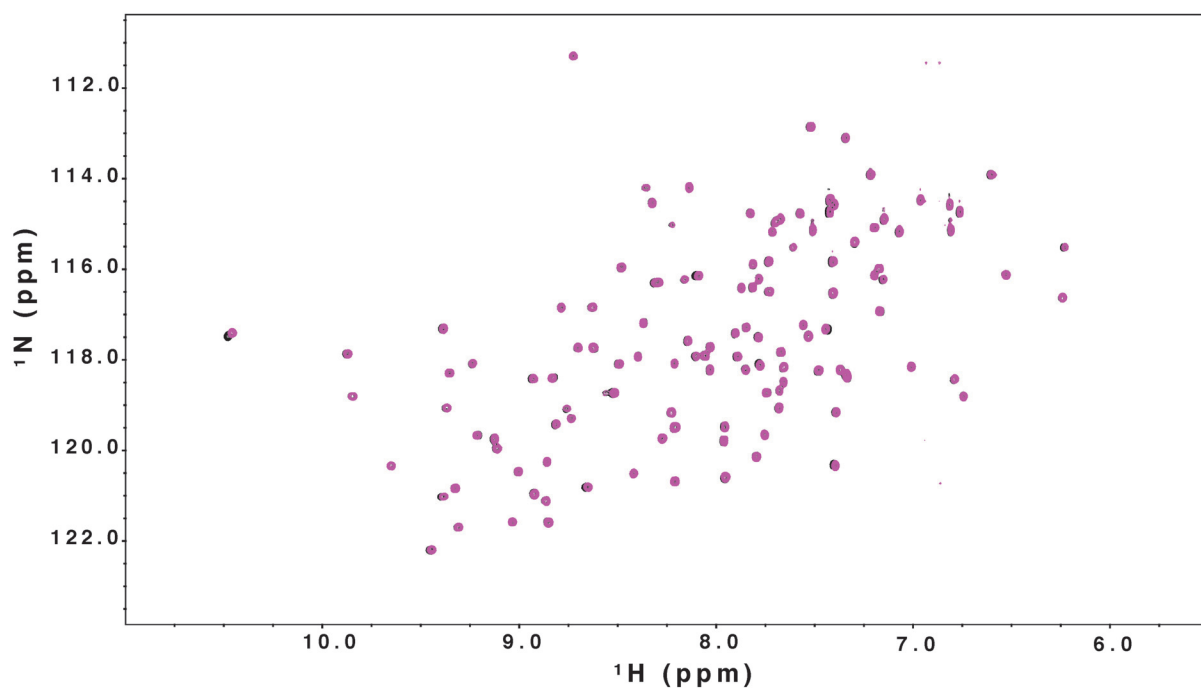


Figure S14. ^{15}N -HSQC spectra of ^{15}N -thioredoxin (Trx). The overlay of spectra recorded without catechol (black spectrum) and after addition of a very large concentration (8mM) of catechol (pink spectrum). The chemical shift are nearly insensitive to large amount of catechol indicating there is no significant interaction with Trx. There was no interaction with thioredoxin reductase (TrxR) demonstrated by absence of saturation transfer difference NMR spectroscopy (STD). TrxR ^{15}N -HSQC but was not fully observable due to large transverse NMR relaxation rates and broad resonance lines.

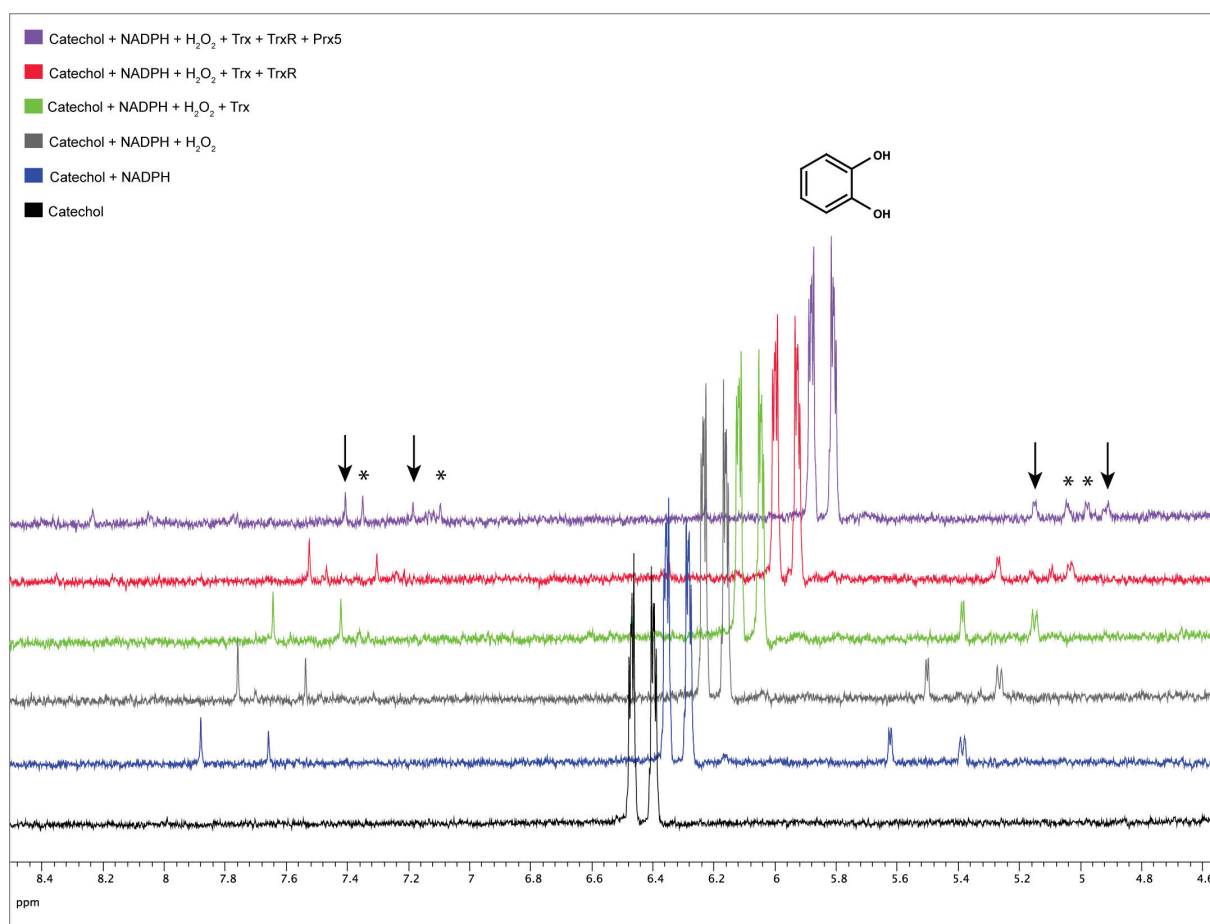


Figure S15. Examining possible side reactions between the catechol derivatives and hPrx5 reaction components (Trx, TrxR, NADPH). A ¹H NMR spectrum was recorded in the presence of catechol (aromatic region) with 600 MHz NMR spectrometer and at 28°C. hPrx5 reaction was broken down sequentially in the addition of the components with catechol present to monitor any chemical transformations. The reaction components were added and incubated for 5 minutes at room temperature followed by a ¹H NMR spectrum. The only chemical reaction monitored was the oxidation of NADPH. The signal of NADPH are indicated with vertical arrows and those of NADP⁺ by an asterisk. NMR assignments were taken from Biological Magnetic Resonance Data Bank entries 55 and 263 for NADPH and NADP⁺, respectively. Overall, the experiments show that catechol is neither transformed in the presence of the different reagents or during the enzymatic reaction.

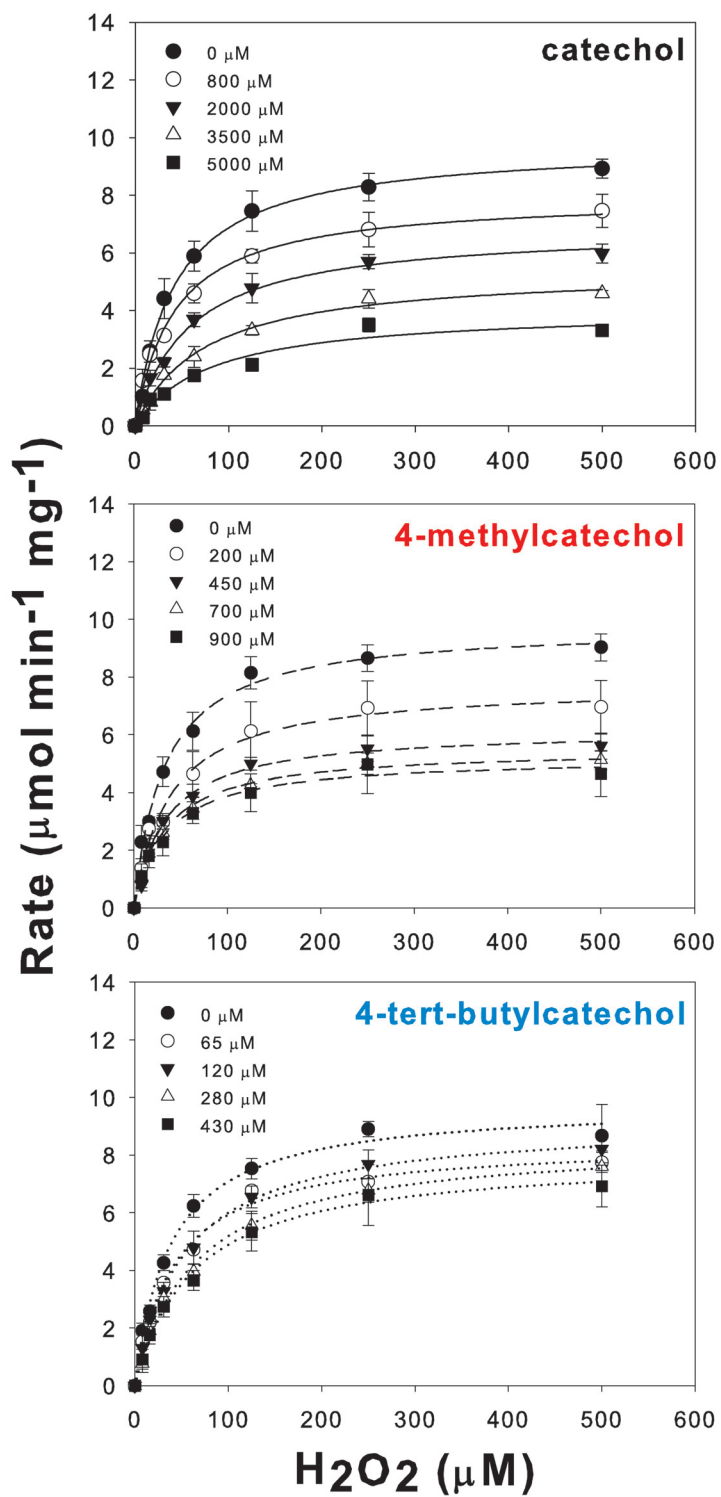
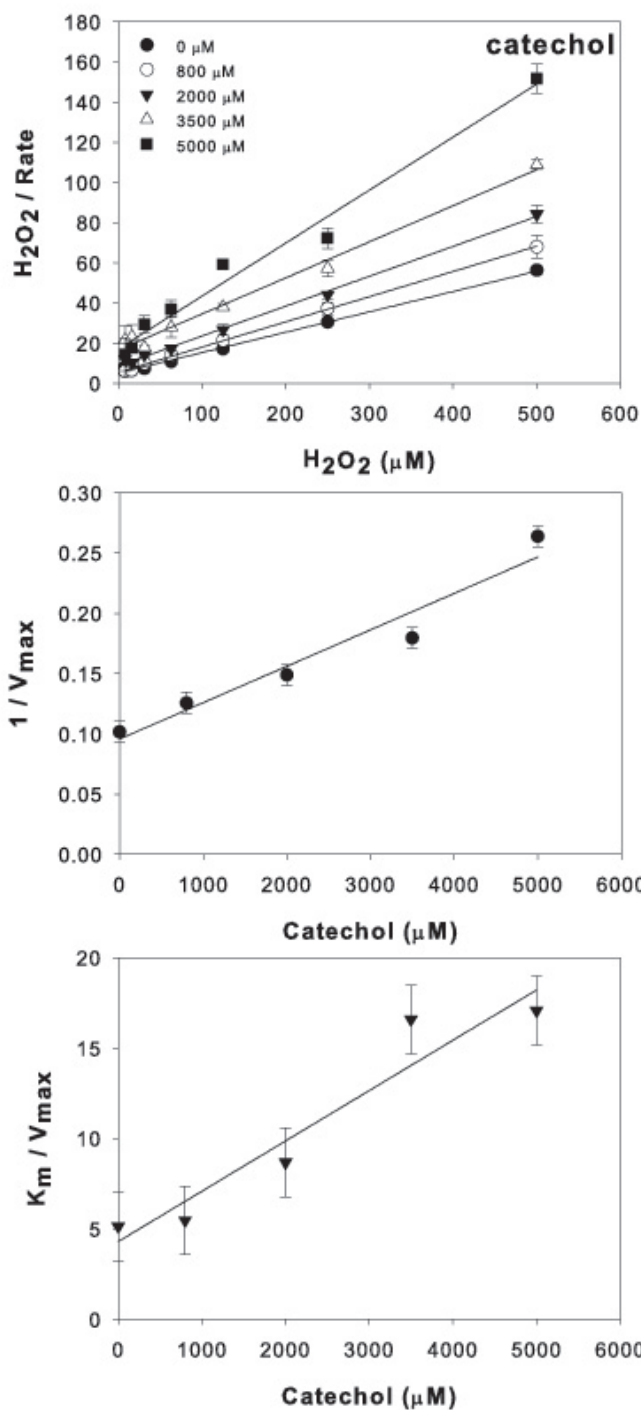


Figure S16: Michaelis-Menten curves for each of the catechol derivatives when monitoring the inhibition mechanism. Plotted here is the initial reaction rates for each catechol (specific concentrations) vs. hydrogen peroxide concentration range. The data was fitted to the Michealis-Menten equation. The changes in the V_{max} and K_m , show that all three catechol derivatives inhibit hPrx5 in a non-competitive mechanism.

(A)



Figures S17 (A-C). Determining the inhibition model by linear regression fitting method. (A) catechol, (B) 4-methylcatechol and (C) 4-*tert*-butylcatechol. Upper plot: the Hanes-Woolf plot with the $[H_2O_2]/\text{initial rate}$ vs. $[H_2O_2]$. The initial rates are expressed in $\mu\text{mol min}^{-1} \text{mg}^{-1}$ in reference to the mg of hPrx5. Both linear fitting and non linear-fitting of analysis generated similar K_m and V_{max} values. Middle plot: The K_i' value was determined from plotting the $1/V_{\text{max}}$ versus the inhibitor concentration. Lower plot: The K_i value was determined from plotting K_m / V_{max} versus the inhibitor concentration. Plotting the data it was not always linear, therefore the fitting was not accurate enough to determine the inhibition model. To further improve the inhibition model discrimination, non-linear model fitting was conducted using DynaFit.

(B)

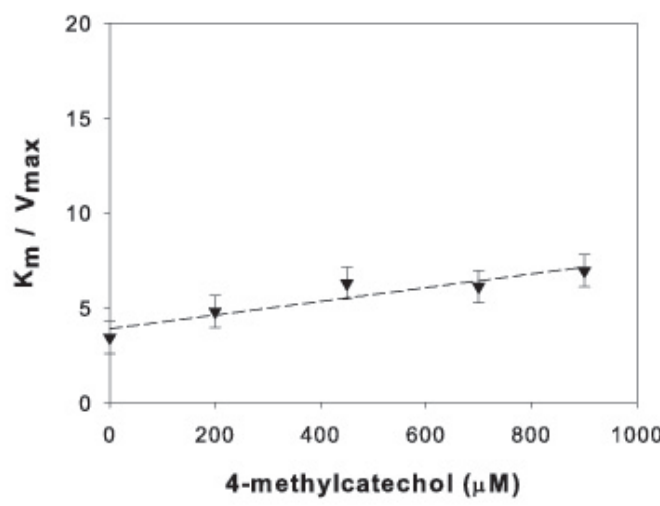
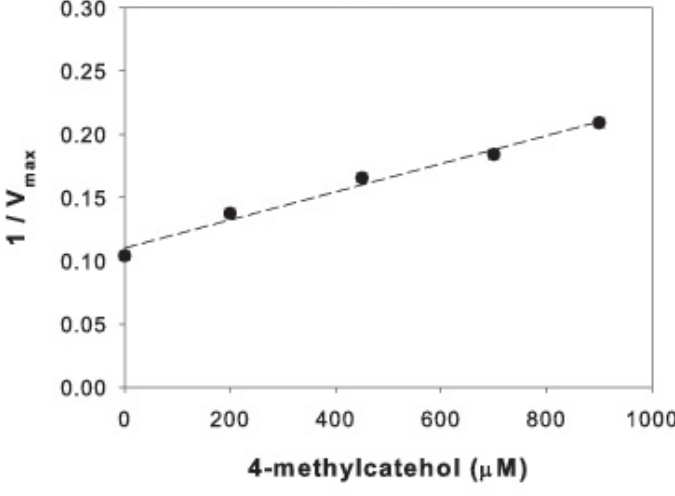
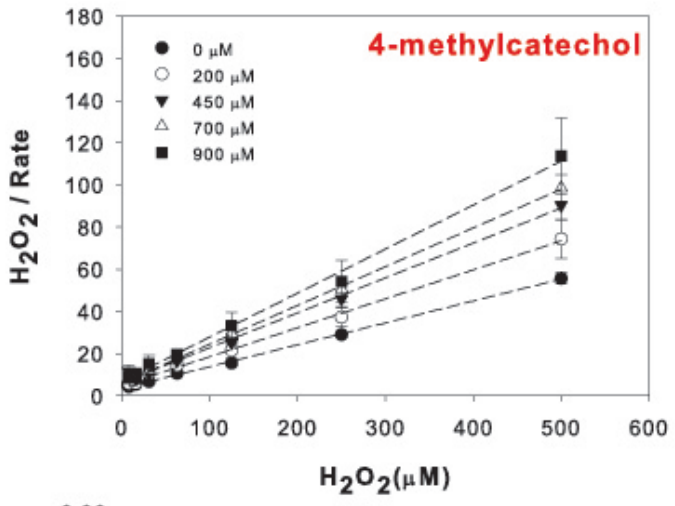


Figure S17 (B).

(C)

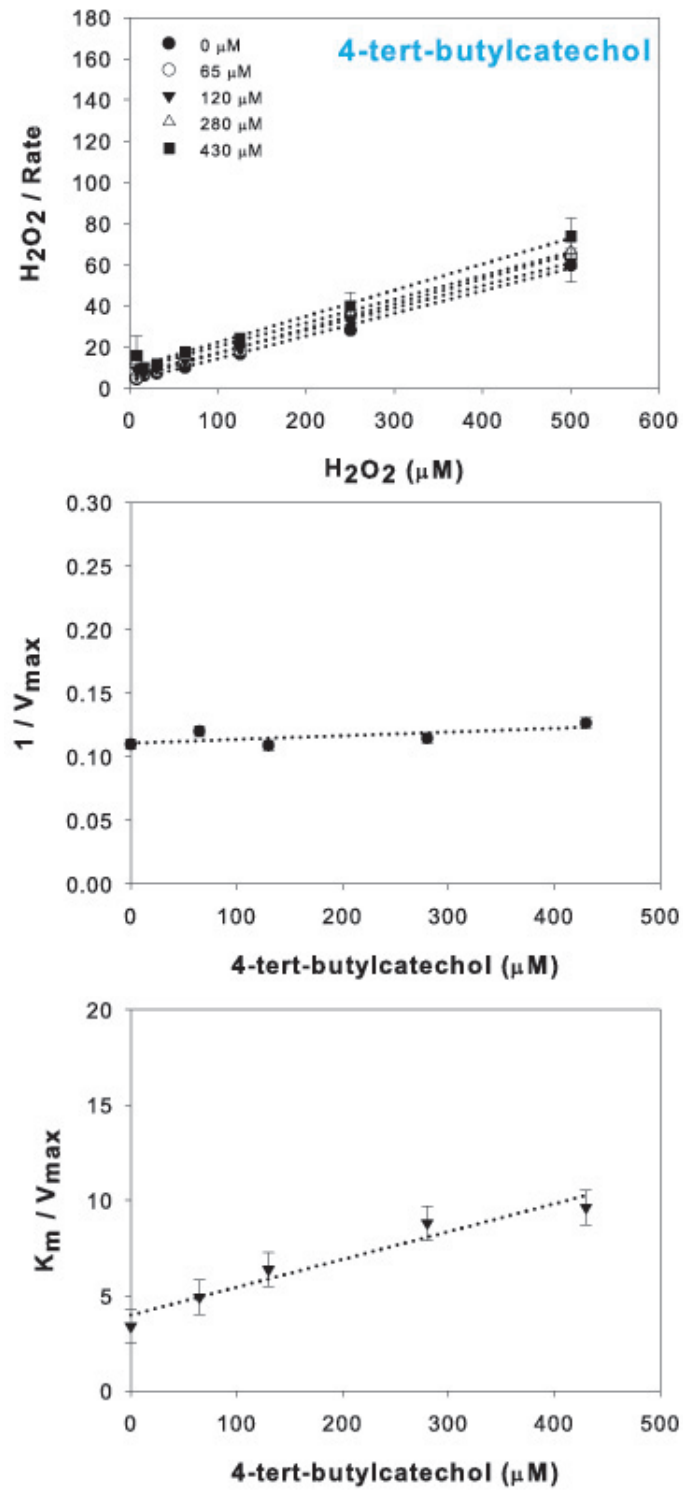


Figure S17 (C).

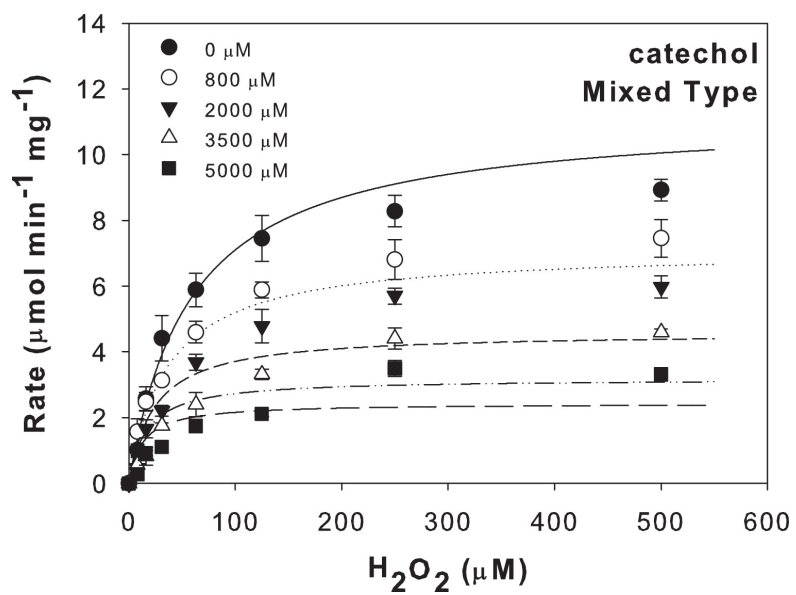


Figure S18. Non-linear inhibition modeling for catechol, fitted by DynaFit to a mixed type non-competitive model. Plotted are the initial rate values of catechol (at specific concentration) vs. the hydrogen peroxide concentration. The data fits poorly for the catechol to a partial mixed type non-competitive model (Fig. 9).

Table S1: Inhibition model analyzed by Michaelis-Menten plots. The K_m (Michaelis-Menten constant) and V_{max} (maximum velocity) values were determined by Michaelis-Menten plots (non-linear) plotting initial rate values vs. $[H_2O_2]$. The K_m values increased and the V_{max} decreased with increasing inhibitor concentrations. This was shown for all three catechol derivatives. Therefore, indicating the catechol derivatives inhibit hPrx5 in a non-competitive manner. The trends of the K_m and V_{max} values complement the Hanes-Woolf plots.

Catechol (μM)	K_m (μM)	V_{max} ($\mu mol\ min^{-1}\ mg^{-1}$)	4-methylcatechol (μM)	K_m (μM)	V_{max} ($\mu mol\ min^{-1}\ mg^{-1}$)	4-tert- butylcatechol (μM)	K_m (μM)	V_{max} ($\mu mol\ min^{-1}\ mg^{-1}$)
0	42.36 \pm 3.33	9.77 \pm 0.22	0	32.96 \pm 2.65	9.77 \pm 0.21	0	38.14 \pm 2.90	9.77 \pm 0.21
800	42.22 \pm 4.07	7.95 \pm 0.22	200	38.35 \pm 3.99	7.72 \pm 0.23	65	42.83 \pm 3.84	8.46 \pm 0.22
200	58.10 \pm 6.95	6.85 \pm 0.26	450	34.29 \pm 4.40	6.16 \pm 0.22	130	53.30 \pm 4.61	9.18 \pm 0.24
3500	74.58 \pm 12.08	5.43 \pm 0.29	700	33.91 \pm 4.87	5.50 \pm 0.22	280	63.91 \pm 6.25	8.49 \pm 0.27
5000	81.45 \pm 18.16	4.05 \pm 0.31	900	34.33 \pm 5.22	5.20 \pm 0.22	430	62.97 \pm 6.54	7.96 \pm 0.26

Table S2: Inhibition model analyzed by Hanes-Woolf plots. The K_m (Michaelis-Menten constant) and V_{max} (maximum velocity) values were determined by Hanes Woolf plots (linear) plotting $[H_2O_2] / \text{rate vs. } [H_2O_2]$. The K_m values increased and the V_{max} decreases with increasing inhibitor concentrations. Therefore, indicating the catechol derivatives inhibit hPrx5 in a non-competitive manner. The trends of the K_m and V_{max} values were complemented by the Michaelis-Menten plots values.

Catechol (μM)	V_{max}		4-methylcatechol		V_{max}		4-tert-butylcatechol		V_{max}	
	K_m (μM)	($\mu\text{mol mg}^{-1} \text{ min}^{-1}$)	(μM)	(μM)	($\mu\text{mol mg}^{-1} \text{ min}^{-1}$)	(μM)	(μM)	($\mu\text{mol mg}^{-1} \text{ min}^{-1}$)	(μM)	($\mu\text{mol mg}^{-1} \text{ min}^{-1}$)
0	50.60 \pm 23.07	9.86 \pm 0.85	0	32.83 \pm 9.41	9.62 \pm 0.36	0	30.81 \pm 9.35	9.11 \pm 0.34		
800	43.51 \pm 18.14	7.96 \pm 0.55	200	34.91 \pm 7.17	7.28 \pm 0.21	65	41.02 \pm 8.89	8.33 \pm 0.28		
200	58.29 \pm 16.17	6.72 \pm 0.39	450	37.93 \pm 6.03	6.04 \pm 0.14	130	58.50 \pm 10.47	9.19 \pm 0.35		
3500	92.35 \pm 15.04	5.57 \pm 0.27	700	33.18 \pm 5.32	5.44 \pm 0.12	280	76.87 \pm 10.60	8.73 \pm 0.31		
5000	64.73 \pm 9.33	3.79 \pm 0.13	900	33.21 \pm 4.69	4.79 \pm 0.09	430	76.01 \pm 9.57	7.90 \pm 0.26		

Nouvelle approche basée sur la dynamique moléculaire et la RMN pour l'étude et l'optimisation de nouveaux inhibiteurs de peroxyrédoxines humaines impliquées dans les chocs ischémiques

La description des interactions protéine-ligand est d'une importance capitale pour la compréhension de phénomènes biologiques de toutes sortes, et dans le procédé de conception de nouvelles molécules bioactives. Avec l'avènement de nouveaux moyens de calculs numériques, la modélisation des équilibres d'interaction utilisant la dynamique moléculaire est de plus en plus utilisée pour l'étude au niveau microscopique permettant ainsi de guider la conception et l'optimisation de nouveaux inhibiteurs biologiques.

Cette thèse est centrée sur l'utilisation de la funnel-métadynamique pour la prédiction d'affinités protéine-ligand et la description des interactions, en corrélation avec des données expérimentales de RMN et de cinétique enzymatique. Sur le modèle biologique des peroxyrédoxines humaines impliquées dans la cascade inflammatoire survenant après un choc ischémique, l'interaction et l'inhibition de composés catéchols ont été étudiées. Les données obtenues par dynamique sur l'interaction de dérivés catéchols ont été utilisées pour guider l'optimisation de cette molécule fragment pour une meilleure affinité et inhibition des peroxyrédoxines humaines. Nos résultats ont montré que la funnel-métadynamique, en plus de permettre la prédiction des affinités protéine-ligand, donnait une description réaliste de l'interaction pouvant mener à l'identification et l'optimisation de nouvelles molécules bioactives dont le potentiel inhibiteur a pu être examiné par cinétique enzymatique. Cette recherche fournit un aperçu des possibilités offertes par les nouvelles méthodes numériques, et leur application en chimie médicinale par exemple.

Mots clés : *Funnel-métadynamique, Interactions, RMN, Catéchol, Peroxyrédoxines humaines, Inhibition*

